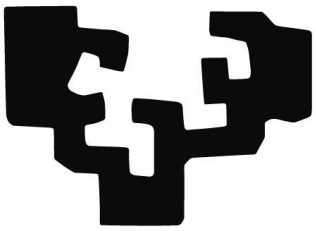
eman ta zabal zazu



Universidad del País Vasco

Euskal Herriko Unibertsitatea

Stochastic Processes for Anomalous Diffusion

Daniel Molina García

2019

Supervisor: Gianni Pagnini

Stochastic Processes for Anomalous Diffusion

Daniel Molina García

2019

Supervisor: Gianni Pagnini

This research is supported by the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323.

Contents

Part I

Initial part: Synthesis

Chapter 1

Introduction

> Mad Hatter: You are trying to understand madness with logic. This is not unlike searching for darkness with a torch.
>
> Brian K. Vaughan, Detective Comics

Anomalous diffusion is a diffusion process which Mean Square Displacement (MSD) is not a linear funtion of time, what is known as normal diffusion. When the relation is faster than linear, it is called superdiffusion and when it is slower, subdiffusion. Physically, the MSD is a measure of the deviation of the position of the particles over time. It can be thought as the amount of space the particles have explored in the system.

On the one hand, anomalous diffusion has been shown to appear extensively in nature and, consecuently, scientists have developed several and different models that can effectively reproduce it. However, the underlying physics of a plethora of experiments is still not well-understood. This is the case of the motion of mRNA molecules inside living E. coli cells [31], where stochastic processes as the continuous-time random walk can explain non-ergodicity [66] but an alternative like the fractional Brownian motion is needed to reproduce p-variation [51]. We find that a family of stochastic processes known as generalised grey Brownian motion [71] can fit correctly that observation. In addition, we also use an explicit time-dependent factor to model non-stationarity, usually referred to aging [96] in the specialized literature. In agreement with our computational results, we were able to obtain analytical expressions for a list of observables too, including temporal and ensemble-average MSD, Ergodicity Breaking parameter, one-point one-time probability density function and p-variation.

On the other hand, many experiments show a characteristic crossover from anomalous to normal diffusion. For example, it happens in viscoelastic systems such as the motion of lipid molecules in lipid bilayer membranes [40, 95]. We studied the motion of particles driven by tempered fractional Gaussian noise which power-law correlations present a cutoff at some mesoscopic time scale. Deriving analytical expressions of the MSD for the overdamped Langevin equation and the fractional Langevin equation, we find that when the truncation is strongly done, the mentioned crossover appears. When the truncation is done by a weak power-law truncation, we also got a different crossover behaviour from faster to slower superdiffusion and from slower to faster subdiffusion.

We were also interested in another process defined through a tempering directly done in the fractional Brownion motion definition. It is known as tempered fractional Brownian motion [62] and, surprisingly, it does not arrive to the same behaviours than our models. Instead, at long times it exhibits localization as Ornstein-Uhlenbeck. We compare both of them. When

the derivative of this tempered fractional Brownian motion is used in the fractional Langevin equation, it leads to ballistic diffusion at long times [18].

In the last part, we were interested in the state of the art in spatial neutral theory of Ecology, which problems sound familiar to experts in Statistical Physics. In fact, ecosystems display a complex spatial organization that ecologists have tried to characterize by observing patterns of biodiversity at different spatial scales. Linking those measures with the causes that originate them is arguably the central problem in Ecology [47]. Ecological neutral theory, which underscores the role of stochastic demographic fluctuations and neglects deterministic effects stemming from fitness differences, has predicted the empirical patterns in communities of competing species. We study non-trivial scaling laws arising at the critical dimension (2D) of spatial neutral models, which are poorly understood. We conclude by discussing models that support non-neutral features.

This thesis by compendium of publications is structured in three parts. The initial part is the Synthesis of the thesis. Chapters and sections of Part I will be described briefly. Chapter 1 is the Introduction where the thesis is presented. In Chapter 2, a Spanish translation of the Introduction is provided. The hypothesis made for this thesis are specified in section 3.1 and section 3.2 enumerates the objectives. Chapter 4 includes the background of the articles that shape this manuscrit and the methodological tools. In more detail, in section 4.1 we highlight some experiments that report anomalous diffusion. If the reader is not familiar with the MSD, it could help read section 4.2.1 before to understand better the details of the applications in this section. In section 4.2 we present the observables of interest for anomalous diffusion. Section 4.3 describes the main stochastic processes for anomalous diffusion, in particular *fractional Brownian motion* and *subdiffusive continuous-time random walk*. Section 4.4 shows how the dependence on the Hurst parameter appears in the MSD of the fractional Brownian motion. A little description of how the process *generalised grey Brownian motion* can be simulated is found in section 4.5. This process will be the subject of interest of the first part of the Results. Section 4.6 presents some models and observables of Community Ecology, the topic of third part of the Results. Methodology is detailed in section 4.7. Results are explicated in three chapters: Chapter 5 is dedicated to the paper "*Fractional kinetics emerging from ergodicity breaking in random media*", Chapter 6 to "*Crossover from anomalous to normal diffusion: truncated power-law noise correlations and applications to dynamics in lipid bilayers*" and Chapter 7 to "*Stochastic spatial models in ecology: a statistical physics approach*". Chapter 8 is focused on the discussion of the Results. Next, the Biblography contains the references of all the citations used in Parts I and II. Part II are the Conclusions. Part III is the Appendix which includes the accepted version of the articles. Full references and links to download the documents are available at the beginning of each article. Finally, at the end of this document, the Acknowledgements section can be found.

Chapter 2

Introducción

> La Afectividad es un estado evolutivo
> superior que no va necesariamente unido
> a la sensibilidad ni a la inteligencia.
>
> Rolando Toro Araneda

Con difusión anómala se hace referencia a procesos de difusión $X(t)$ en los cuales el desplazamiento cuadrático medio (MSD) $\langle X^2(t)\rangle$ no es una función lineal de la variable tiempo t (lo que se conoce como difusión normal). Cuando la relación es más rápida que lineal, se le llama superdifusión, y cuando es más lenta, subdifusión. Físicamente, el MSD es una medida de las desviación de la posición de las partículas con el tiempo. Se puede imaginar como la cantidad de espacio que las partículas han explorado en el sistema.

La difusión anómala aparece constantemente en la naturaleza y por ello los científicos han desarrollado diferentes modelos que pueden reproducirla efectivamente. Sin embargo, la física subyacente de una gran cantidad de experimentos aun no se comprende correctamente. Este es el caso del movimiento de las moléculas de ARN mensajero dentro de bacterias E. coli, donde los más importantes procesos estocásticos fallan al intentar explicar todas sus características al mismo tiempo. Por ejemplo, el caminante aleatorio con tiempo continuo (CTRW) puede explicar la falta de ergodicidad pero no la variación-p. Por otro lado, el movimiento Browniano fraccionario (fBm) puede explicar su variación-p pero se trata de un proceso ergódico. Hemos analizado una clase de procesos estocásticos conocida como movimiento Browniano gris generalizado (ggBm) que además de reproducir difusión anómala, no es ergódico y tiene una variación-p como la del fBm, siendo por tanto un buen candidato para explicar el experimento citado anteriormente. Además, hemos incluído un factor con una dependencia temporal explícita para que sea un proceso no estacionario, característica que se suele denominar envejecimiento. De forma coincidente con los resultados computacionales, hemos podido encontrar expresiones matemáticas para muchos observables incluyendo el MSD promediado colectivamente y temporalmente, el parámetro de Rotura de la Ergodicidad, la función densidad de probabilidad en un punto y en un tiempo y la variación-p.

Por otro lado, muchos experimentos muestran una transición característica de un régimen de difusión anómala a otro de difusión normal. Por ejemplo, esto ocurre en sistemas viscoelásticos como el movimiento de moléculas de lípidos en membranas bicapa de lípidos. La segunda parte de esta tesis está dedicada al estudio de procesos estocásticos que mediante una truncamiento de la función de autocorrelación del ruido o fuerza estocástica en la ecuación de Langevin sobredimensionada o la ecuación de Langevin fraccionaria sobredimensionada, se puede conseguir este tipo de transición. Esto ocurre cuando el truncamiento se hace de forma exponencial o mediante una ley de potencias suficientemente fuerte. Si la ley de potencias es débil, se obtienen transiciones de un régimen de rápida superdifusión a otra más

lenta, y de un régimen de subdifusión lenta a otra más rápida, respectivamente. En esta parte también se consideran otros procesos en los que el truncamiento se realiza directamente sobre la definición del fBm original de Mandelbrot y Van Ness. Se conoce como movimiento Browniano fraccionario templado (tfBm) y, sorprendentemente, no tiene las mismas propiedades que los anteriores modelos. En su lugar, a tiempos largos presenta localización como el proceso de Ornstein-Uhlenbeck. Hemos comparado ambos procesos. Cuando la derivada del tfBm es usada en ecuación de Langevin fraccionaria, se obtiene difusión balística para tiempos largos.

En la última parte estamos interesados en los últimos avances en teoría neutral espacial del campo de la Ecología, cuyos problemas resultan familiares a los expertos en Física Estadística. De hecho, los ecosistemas reproducen una organización espacial compleja que los ecólogos han intentando caracterizar observando los diferentes patrones de biodiversidad a distintas escalas espaciales. Relacionar estas medidas con las causas que las originan es probablemente el problema central de la Ecología. La teoría neutral de la Ecología, que subraya el papel de las fluctuaciones demográficas estocásticas y rechaza los efectos deterministas procedentes de diferencias adaptativas, ha predicho los patrones empíricos en comunidades de especies que compiten entre si. Hemos estudiado las leyes de escalado que surgen en la dimensión crítica (2D) de los modelos espaciales neutrales, los cuales no se comprenden del todo a día de hoy. Acabamos discutiendo algunos modelos con características no neutrales.

Chapter 3

Hypothesis and objectives

> They are proud in humility, proud in that they are not proud.
>
> Robert Burton,
> The Anatomy of Melancholy

3.1 Hypothesis

For the first article, the following stochastic model is considered

$$X(t) = l_\beta X_{\text{gen}}, \tag{3.1}$$

where l_β is a random variable distributed according to a distribution depending on the parameter β and X_{gen} is some ergodic Gaussian process.

The interpretation of this model is fixed in that article, and it is the following one:

- The stochastic process is the consecuence of a specific interaction of a medium with the particles.
- In a homogeneous medium, the process describing the particles would be the same for all of them: X_{gen}.
- However, the considered medium is the cause that results in the position of each particle being multiplied by a constant determined through the random variable l_β (the medium associates a "length scale" to each particle).
  - In that sense, it is said that the medium is heterogeneous.
  - It must not be confused with other heterogeneous kind of media.
  - In our case, the length scale associated to each particle does not change with time nor position. The only exception will be the extension used to reproduce aging, which will consider a prefactor dependent on time.

In the third article, some ecological models are stated to be neutral. Neutral Ecology assumes that the differences between individuals of an ecological community of trophically similar species are irrelevant to their success. This implies that biodiversity is originated randomly, underscoring the role of stochastic demographic fluctuations. This hypothesis has elicited heated controversies because classical ecological concepts, e.g. niches, might be irrelevant.

Spatial models are also focused. The main characteristic of these models is that they control explicitly the distances between individuals. It is usually done by considering that they are placed in a lattice structure.

3.2 Objectives

This thesis is based on three scientific articles. The main objective of the thesis is to present the scientific results that gave rise to the manuscripts. With respect to the first article, the following objectives are addressed:

- Have knowledge about the main properties of the stochastic process known as generalised grey Brownian motion (ggBm).
- Find candidate applications that could fit the properties of the ggBm.
- Calculate numerically (and analytically if possible) the known observables of the chosen application for the ggBm.
- Determine if the ggBm properties fit well the data of the application (or even better than other stochastic processes).

For the second article:

- Start by showing the dynamics of a stochastic particle driven by fractional Gaussian noise and Langevin equation or generalised Langevin equation.
- Calculate what happens when the noise is tempered by:
  - An exponential cutoff, or
  - A power-law cutoff.
- Study the process tempered fractional Brownian motion to see the similitudes and differences with the results of above.

For the third article:

- Discuss the state of the art in spatial neutral theory.
- Describe the ecological patterns that produce the voter model with speciation by presenting mathematical and computational results.
- Consider also non-neutral models and highlight their differences with respect to neutral models.

Chapter 4

Background

> I believe, indeed, that overemphasis on the purely intellectual attitude, often directed solely to the practical and factual, in our education, has led directly to the impairment of ethical values.
>
> Albert Einstein,
> The Need for Ethical Culture

4.1 Presentation of applications depicting anomalous diffusion

Next, we describe some experimental results that illustrate the presence of anomalous diffusion in the real world. Sometimes, it is not easy to distinguish which kind of MSD is considered when reading the original papers. Because of that, expressions and notations of the original papers for the MSD will be preserved to avoid any risk of misinterpretation and, as a consequence, the notation will not be consistent with the rest of this manuscript.

4.1.1 Motion of mRNA molecules in E. coli cells

The motion of individual molecules inside E. coli cells has been examined by tracking fluorescently messenger RNA which was free to move in the cytoplasm [31]. During 30 minutes, cells were tracked at 1 frame/sec and the mRNA moved spanning all the cell multiple times. It was observed that the motion was discontinuous, alternating periods of almost localizated motion with fast jumps to new positions. In addition, it was found the following law for the motion of the mRNA [31]

$$\left\langle \delta^2(\tau) \right\rangle \sim \Gamma \tau^{\alpha}, \qquad \alpha < 1, \tag{4.1}$$

where

$$\delta = |\boldsymbol{r}(t+\tau) - \boldsymbol{r}(t)| \tag{4.2}$$

is the particle displacement between two time points elapsed by τ. While the generalised diffusion coeffient Γ changed considerably between cells and experimental conditions, the subsiffusion exponent α had almost no dependence on particle size, growth conditions and genetic background (e.g. cytoskeletal mutants). Averaging over 21 trajectories, they found a value of $\alpha = 0.70 \pm 0.07$.

4.1.2 Infection pathway of Adeno-Associated virus

The tracking of single molecules in real time has allowed to visualize the infection pathway of single viruses in living cells. In particular, it has been studied the motion of adeno-associated viruses (AAV) into living HeLa cells [86]. After membrane penetration, endosome diffusion was observed with only one virus per endosome. Direct motion and normal and anomalous difussion of the endosome with the virus are reported in the cytoplasm and the nucleus. In the cytoplasm, 51 of 113 viral particles showed anomalous diffusion, following a

$$\left\langle r^2 \right\rangle = 4Dt^\alpha \tag{4.3}$$

law with D between 0.3 and 1.5 $\mu m^2/s$ and α in the range of 0.5 to 0.9. In the nucleus, 23 of the trajectories (total is said to be more than 100) again depicted anomalous diffusion, in this case somewhat slower, with $D \in [0.1, 0.5]$ $\mu m^2/s$ and $0.6 < \alpha < 0.9$.

4.1.3 Motion of lipid granules in yeast cells

The viscoelastic characteristics of the cytoplasm of fission yeast *Schizosaccharomyces pombe* cells were studied by analysing the motion of lipid granules [93]. The granules are embedded in a cytoskeleton made of a sparse microtubule network and actin filaments. Polymer networks of the cytoskeleton may limit the fluctuation of a granule position. To clasify the motion, time series of its position $\vec{r}(t) = (x(t), y(t))$ were recorded and calculated the mean squared displacement, $\left\langle |\Delta\vec{r}(t)|^2 \right\rangle$, where t is the time lag. The average was taken within a single trajectory over time. Authors fitted the results to

$$\left\langle |\Delta\vec{r}(t)|^2 \right\rangle \propto t^\alpha \tag{4.4}$$

and they obtained $\alpha = 0.737 \pm 0.003$ for $N = 266$ granules inside living cells. They also performed experiments in which actin filaments were previously disrupted. Before disruption, they found $\alpha = 0.734 \pm 0.004$ ($N = 59$, a subset of the previous granules) and, after treatment, $\alpha = 0.755 \pm 0.006$ ($N = 52$).

4.1.4 Diffusion of telomeres in the nucleus of mammalian cells

Experiments performed in living human cells, study the diffusion properties of telomeres in the nucleus from 10^{-2} to 10^4 seconds [10]. Typically, about 60 telomeres were observed in each cell. The MSD was calculated for each telomere by finding the average displacement between each two time points. Time interval was $\Delta = n\tau$, being τ the measurement time interval and n an integer. The averaging was done over all the measured time according to the formula

$$\langle r^2(\Delta) \rangle = \frac{1}{N-n} \sum_{m=1}^{N-n} [\boldsymbol{r}((m-1)\tau + \Delta) - \boldsymbol{r}((m-1)\tau)]^2, \tag{4.5}$$

being N the number of measured points and $\boldsymbol{r}$ the position vector of the particle. It was found that diffusion was subdiffusive

$$\langle r^2(t) \rangle = At^\alpha, \qquad \alpha < 1, \tag{4.6}$$

at short time scales of 10^{-2} to 10^2 seconds, then a transient and finally changed to normal diffusion at the time range of 10^2 to 10^4 seconds. 419 telomeres ware analyzed in the range 10^{-2} to 1 second and the values found for the coefficient A varied considerably from one telomere to another. In contrast, α changed much less, being $\alpha = 0.32 \pm 0.12$. In the range 10^0 to $2 \cdot 10^2$ seconds, it was found $\alpha = 0.51 \pm 0.20$ using 151 telomeres.

4.2 Presentation of the observables

Next, we present the main observables of interest in anomalous diffusion in this thesis. As it is usually done, we will work in one dimension.

4.2.1 Mean Square Displacement

In this thesis we distinguish between two types of Mean Square Displacements (MSD). The EAMSD and the TAMSD.

Let us consider the stochastic process $X(t)$. If $X(t)$ has a known one-point one-time probability density function $P(x,t)$, the Ensemble-Average Mean Square Displacement (EAMSD) is equal to the second moment, which can be calculated as

$$\left\langle X^2(t) \right\rangle = \int_{-\infty}^{\infty} x^2 P(x,t). \tag{4.7}$$

As opposed to the EAMSD, it is possible to consider the Time-Average Mean Square Displacement (TAMSD) of a single trajectory

$$\overline{\delta^2(\Delta)} = \frac{1}{T-\Delta} \int_0^{T-\Delta} \left[X(\tau+\Delta) - X(\tau)\right]^2 d\tau, \tag{4.8}$$

where T is known as the measurement time and Δ the lag time. To obtain a unique result also at finite measurement times, an additional ensamble average can be applied to the time average, which is known as EATAMSD

$$\left\langle \overline{\delta^2(\Delta)} \right\rangle = \frac{1}{T-\Delta} \int_0^{T-\Delta} \left\langle \left[X(\tau+\Delta) - X(\tau)\right]^2 \right\rangle d\tau. \tag{4.9}$$

In practice, when dealing with discrete trajectories, e.g. by doing simulations or managing experimental trajectories, the quantity $\delta^2(\tau,\Delta) = [X(\tau+\Delta) - X(\tau)]^2$ is used for the average. To obtain the EAMSD, using $\tau = 0$ and $\Delta = t$, $\delta^2(0,t)$ is averaged over N trajectories $X_i(t)$,

$$\left\langle X^2(t) \right\rangle_N = \frac{1}{N} \sum_{i=1}^{N} [X_i(t) - X_i(0)]^2. \tag{4.10}$$

To calculate the TAMSD of one trajectory defined in discrete times from 0 to T, it is enough to fix Δ and average $\delta^2(\tau,\Delta)$ over M values of τ_j from 0 to $T-\Delta$. It can be simply written

as

$$\overline{\delta^2_M(\Delta)} = \frac{1}{M}\sum_{j=1}^{M}[X(\tau_j+\Delta)-X(\tau_j)]^2. \tag{4.11}$$

As it was said before, it is common to average TAMSD also between trajectories, what is known as EATAMSD. It results

$$\left\langle\overline{\delta^2_M(\Delta)}\right\rangle_N = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}[X_i(\tau_j+\Delta)-X_i(\tau_j)]^2. \tag{4.12}$$

4.2.2 p-variation

Let $X(t)$ be a stochastic process observed on the time interval $[0,T]$. Let us divide $[0,T]$ in 2^n identical subintervals. The 2^n+1 limits of the subintervals are $t_k = T\,k/2^n, k \in \{0,1,\cdots,2^n\}$. For these points it is posible to calculate the so called sample p-variation [52]

$$V_n^{(p)}(t_k) := \sum_{i=1}^{k}|X(t_i)-X(t_{i-1})|^p. \tag{4.13}$$

Sample p-variation can be calculated for every $t \in [0,T]$ using interpolation between t_k, even if it is not usually done in practice. When the process is defined for every t, the following definition that takes care of all the details is frequently used

$$V_n^{(p)}(t) = \sum_{j=0}^{2^n-1}\left|X\left(\frac{(j+1)T}{2^n}\wedge t\right)-X\left(\frac{jT}{2^n}\wedge t\right)\right|^p, \tag{4.14}$$

where $a\wedge b = \min\{a,b\}$. That is, it uses definition (4.13) for the biggest $t_k \le t$ and then it adds $|X(t)-X(t_k)|^p$ (the rest of the summands are equal to zero).

p-variation is defined as

$$V^{(p)}(t) = \lim_{n\to\infty} V_n^{(p)}(t). \tag{4.15}$$

For large enough n, sample p-variation is a good approximation of p-variation.

p-variation is a generalization of the concept of total variation of a function ($p=1$) or the quadratic variation ($p=2$). p-variation can be understood as a measure of the fluctuations of the process and then, also of its fractal dimension and the Hurst parameter H [52].

There is an important lemma that establishes [51]: Let $X(t)$ be continuous. If for some $p_*>0$, $V^{(p_*)}(t)$ is positive and finite (let us call it $f(t)$), then p-variation satisfies

$$V^{(p)}(t) = \begin{cases}\infty, & p<p_*,\\ f(t), & p=p_*,\\ 0, & p>p_*.\end{cases} \tag{4.16}$$

p-variation test

In article [53] it is presented a test based on the previous lemma to distinguish between stochastic processes given sampled trajectories. Using previous notation to simplify the explanation, the idea is to calculate the finite sums $V_n^{(p_*)}(t)$ for the p_* of the considered stochastic processes (in particular, they use the $p_* = 1/H$ of fractional Brownian motion and the $p_* = 2$ of subdiffusive continuous-time random walk) and see what happens when increasing n. If $V_n^{(p_*)}(t)$ converges to 0 or diverges when increasing n, it will be evident that the trajectory does not come from the stochastic process associated to p_*, so it can be discarded. If $V_n^{(p_*)}(t)$ converges to a curve when increasing n, and it is the $f(t)$ of the stochastic process associated to p_*, there are evidences to assert that the trajectory was originated by that process.

You can find the graphs of the application of this test to fractional Brownian motion in section 4.3.2 and to subdiffusive continuous-time random walk in section 4.3.3.

4.2.3 Ergodicity Breaking

In dynamical systems or Markov chains, the state space and dynamics are known and it can be established whether the *system* is ergodic or not. In the afirmative case, any non-pathological observable would be ergodic [63]. In previous reference it is indicated that for stochastic processes, ergodicity has to be defined for a *particular observable* y by checking the equivalence of its ensemble and time average:

$$\langle y \rangle \stackrel{?}{=} \lim_{T\to\infty} \frac{1}{T} \int_0^T y(t')dt'. \tag{4.17}$$

There are different reasons such that the equivalence cannot be fulfilled. In particular, the temporal-average limit may not exist. When the state space of the system is made of disconnected parts and the time-average differs depending on the starting point, it is called *strong ergodicity breaking* [63]. When the state space is connected but a single trajectory cannot sample it fully because the motion is "freezed", it is termed *weak ergodicity breaking* [63].

Having said that, in practice, the more common observable for which ergodicity is checked is the MSD and it is usually said that the stochastic process is ergodic when it is ergodic for the MSD. We will consider the following definition for ergodicity:

> If the EATAMSD $\langle\overline{\delta^2}(\Delta)\rangle$ is equal to the EAMSD $\langle X(\Delta)\rangle$, and if the variance of the TAMSD tends to zero when T is long, the process is *ergodic* (the distribution of the TAMSD tends to a delta function centered on the EAMSD) [26].

A dimensionless measure of Ergodicity Breaking is the Ergodicity Breaking parameter

$$E_B = \frac{\mathrm{Var}\left\{\overline{\delta^2}\right\}}{\left\langle\overline{\delta^2}\right\rangle^2} \tag{4.18}$$

$$= \frac{\left\langle\left[\overline{\delta^2}\right]^2\right\rangle}{\left\langle\overline{\delta^2}\right\rangle^2} - 1. \tag{4.19}$$

As a consequence of the previous condition for ergodicity, an ergodic process satisfies $E_B = 0$ when T is long.

4.2.4 Aging

For a stationary process, the two-times correlation functions are only function of the time difference

$$\langle X(t)X(s)\rangle = f(|t-s|). \tag{4.20}$$

However, there are other systems (spin glasses, coloidal glasses, gels, turbulent systems, etc.) in which stationary properties are violated. For example, they may be characterized by correlation functions of the form

$$\langle X(t)X(s)\rangle = f(t/s), \qquad s \geq t. \tag{4.21}$$

The origin of times cannot be arbitrary in these cases. A generalization of the Wiener-Khinchin theorem has been developed for this kind of aging systems [11] and, in particular, it has been derived the condition for ergodicity in terms of the correlation function of a given observable.

Aging can be formally defined as the explicit dependence of observables on the time difference between the original preparation of the system and the start of the recording data [84]. This time difference is usually known as the aging time t_a.

In some experiments, the aging time can be set deliberately. However, in other systems, it cannot be precisely measured. In both cases it is important to know how it affects the measures of the observables.

In the case of aging, the EAMSD of a process X is formulated in the form

$$\left\langle X^2(t)\right\rangle_a = \left\langle [X(t+t_a) - X(t_a)]^2 \right\rangle. \tag{4.22}$$

On the other hand, the TAMSD is calculated as

$$\overline{\delta_a^2(\Delta)} = \frac{1}{T-\Delta}\int_{t_a}^{t_a+T-\Delta} [X(\tau+\Delta) - X(\tau)]^2 d\tau. \tag{4.23}$$

4.2.5 Determining the correct model from different observables

As we saw in section 4.2.2, p-variation test can be a useful technique to distinguish between models given a single trajectory. Tests for ergodicity have been applied [42, 49, 91] too. In reference [64], it is presented a flow-chart to decide the cause of subdiffusion. The first step is to check if the process is ergodic: If it is not ergodic, the process is associated to trapping (continuous-time random walk). Else, the second step is to distinguish between percolation-like structures (random walk on a percolation cluster) or viscoelastic media (fractional Brownian motion). To achieve this, they study (in two dimensions) the space-filling properties of the trajectories, considering in particular the walk dimension and the fractal dimension. Reference [63] is dedicated to review the main statistical tools available to reconstruct the underlying physics from the details of the dynamics. It considers tests for the following characteristics:

- Evaluation of the subdiffusive exponent
- Stationarity
- Ergodicity
- Fundamental moment and p-variation
- Gaussianity
- Correlations
- Geometry

4.3 Presentation of stochastic processes for anomalous diffusion

In the next sections, some important stochastic processes will be described. Main results that are relevant for the understanding of the objectives of this thesis will be presented and plotted with simulation results. An extensive overview over different popular anomalous diffusion models is reference [66].

4.3.1 Classical diffusion

Robert Brown reported in 1828 how small granules made from pollen grains was moving randomly when inmersed in water. Einstein, in 1906 published a paper in which it was explained in detail how the motion of those grains was the result of the thermal movement of individual water molecules. This explanation was an evidence that atoms and molecules exist, which would be later verified by Jean Perrin in 1908. Einstein derived a diffusion equation for the one-time one-point probability density function $P(x,t)$ of the diffusing granules

$$\frac{\partial}{\partial t}P(x,t) = \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}P(x,t), \tag{4.24}$$

which is equivalent to the Fick's second law. The solution is the Gaussian probability density function (pdf)

$$P(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp^{-\frac{x^2}{2\sigma^2 t}}. \tag{4.25}$$

This process is known, in particular, as Brownian motion (Bm) $B(t)$. The EAMSD is

$$\left\langle B^2(t)\right\rangle = \sigma^2 t, \tag{4.26}$$

and the TAMSD [65]

$$\overline{\delta^2(\Delta)} = \sigma^2\Delta, \qquad \Delta \ll T. \tag{4.27}$$

Also, without the need of very large T, the EATAMSD is

$$\left\langle\overline{\delta^2(\Delta)}\right\rangle = \sigma^2\Delta. \tag{4.28}$$

Bm is then ergodic. These properties (Gaussian pdf and EAMSD and TAMSD being linear in time) are the classical properties of diffusion. In next sections, other models that does not

fulfill these conditions will be introduced to answer the necessity of new types of diffusion. They are known as anomalous diffusion. p-variation of Bm is

$$V^{(p)}(t) = \begin{cases} +\infty & \text{if } p < 2, \\ \sigma^2 t & \text{if } p = 2, \\ 0 & \text{if } p > 2. \end{cases} \tag{4.29}$$

Bm can be expressed as the Langevin equation

$$\frac{dB(t)}{dt} = B'(t) \tag{4.30}$$

where $B'(t)$ is white Gaussian noise with zero mean and autocorrelation function

$$\left\langle B'(t)B'(s)\right\rangle = \sigma^2\delta(t-s). \tag{4.31}$$

4.3.2 Fractional Brownian motion

In the original paper by Mandelbrot and Van Ness [58], the fractional Brownian motion (fBm) $B_H(t)$ is defined. It is a real random function where $0 < H < 1$ is the Hurst index. A similar definition of fBm, more friendly for physicists, is the following

$$B_H(t) := \frac{1}{\Gamma(H+1/2)}\Bigg\{ \int_{-\infty}^{0} \left[(t-t')^{H-1/2} - (-t')^{H-1/2}\right] B'(t')dt' + \int_0^t (t-t')^{H-1/2} B'(t')dt' \Bigg\}, \quad t > 0, \tag{4.32}$$

where $B'(t)$ is white Gaussian noise. It can be simply written in the alternative way

$$B_H(t) = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{\infty} \left[(t-t')_+^{H-1/2} - (-t')_+^{H-1/2}\right] B'(t')dt', \tag{4.33}$$

by using operator $(x)_+ = \max\{0, x\}$. $B_{1/2}(t)$ is the Brownian motion $B(t)$. $B_H(t)$ satisfies

$$\langle B_H(t)\rangle = 0. \tag{4.34}$$

The covariance function (see Figure 4.1) is

$$\gamma_H(t,s) = \frac{\sigma^2 V_H}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \tag{4.35}$$

where

$$V_H = \frac{1}{\Gamma(2H+1)\sin(\pi H)}. \tag{4.36}$$

The variance or EAMSD of fBm is

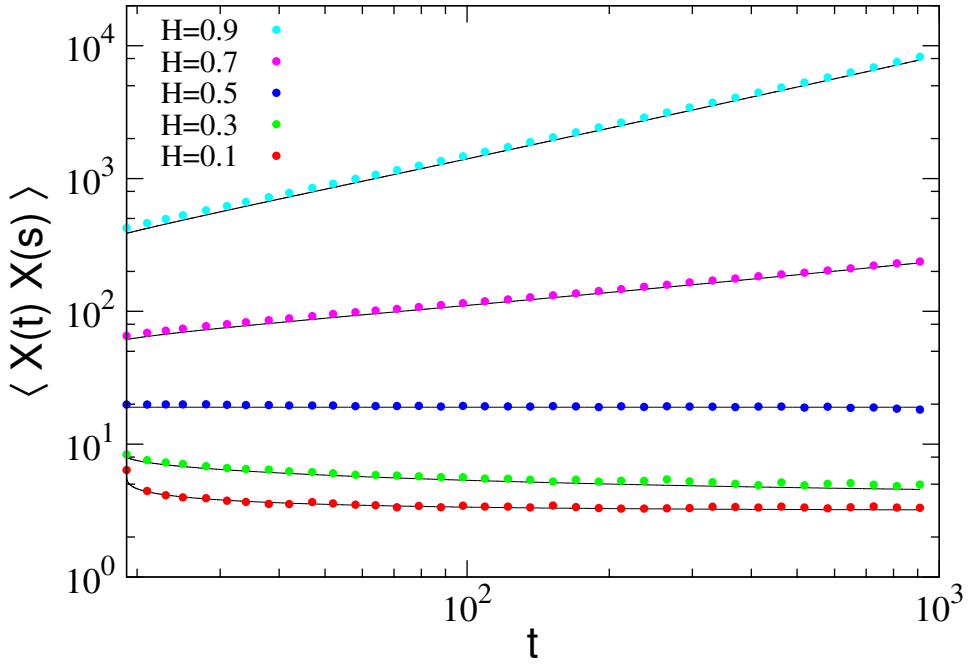


Figure 4.1: Covariance of fBm for different values of H and $s = 19$. Points are simulation data and black lines the theoretical values (4.35).

$$\langle B_H^2(t)\rangle = \sigma^2 V_H t^{2H}. \tag{4.37}$$

Note that in the literature it is common to consider a "normalized" fBm in which the term V_H does not appear. See section 4.4 for a derivation of the EAMSD and see how V_H emerges from Mandelbrot's definition of fBm. In Figure 4.2 the EAMSD of fBm is shown. $B_H(t)$ is a Gaussian process, so $B_H(t) \sim N(0, \sigma^2 V_H t^{2H})$ and it has one-point one-time pdf

$$P(x,t) = \frac{1}{\sqrt{2\pi\sigma^2 V_H t^{2H}}} e^{-\frac{x^2}{2\sigma^2 V_H t^{2H}}}. \tag{4.38}$$

Its pdf can be observed in the right plot of Figure 4.3. The diffusion equation of fBm is

$$\frac{\partial}{\partial t}P(x,t) = \sigma^2 V_H H t^{2H-1} \frac{\partial^2}{\partial x^2} P(x,t). \tag{4.39}$$

The EATAMSD is [26]

$$\left\langle \overline{\delta^2(\Delta)} \right\rangle = \sigma^2 V_H \Delta^{2H}. \tag{4.40}$$

For long measurement times, TAMSD is self-averaging and it holds [66]

$$\overline{\delta^2(\Delta)} = \sigma^2 V_H \Delta^{2H}, \quad T \gg \Delta. \tag{4.41}$$

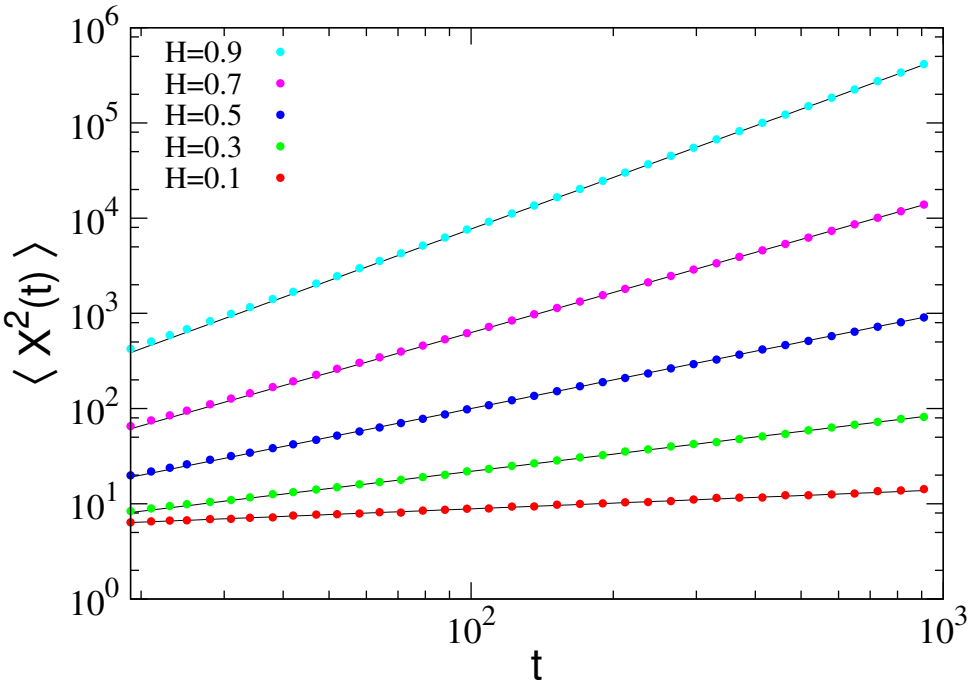


Figure 4.2: Variance of fBm for different values of H. Points are simulation data and black lines are the theoretical values (4.37).

In Figure 4.4 it can be seen how TAMSD of different trajectories surround the EATAMSD, and how they are almost equal for small values of Δ. Notice that variance of TAMSD is large when $\Delta \approx T$. The Ergodicity Breaking parameter is 0 becasue of the process being ergodic. The asymptotic behaviour of $E_B(T)$ for fBm is [26]

$$E_B(T) \sim \begin{cases} k(H)\frac{\Delta}{T}, & 0 < H < 3/4, \\ k(H)\frac{\Delta}{T}\ln(T), & H = 3/4, \\ k(H)\left(\frac{\Delta}{T}\right)^{4-4H}, & 3/4 < H < 1, \end{cases} \tag{4.42}$$

where

$$k(H) = \begin{cases} \int_0^\infty ((\tau+1)^{2H} + |\tau-1|^{2H} - 2\tau^{2H})^2 d\tau, & 0 < H < 3/4, \\ 9/16, & H = 3/4, \\ \left(\frac{4}{4H-3} - \frac{4}{4H-2}\right) H^2(2H-1)^2, & 3/4 < H < 1. \end{cases} \tag{4.43}$$

In Figure 4.5, $E_B(T)$ (4.42) is represented for some values of H along some simulation results. p-variation of fBm has the form [53]

$$V^{(p)}(t) = \begin{cases} +\infty & \text{if } p < 1/H \\ k\,t & \text{if } p = 1/H \\ 0 & \text{if } p > 1/H \end{cases}, \tag{4.44}$$

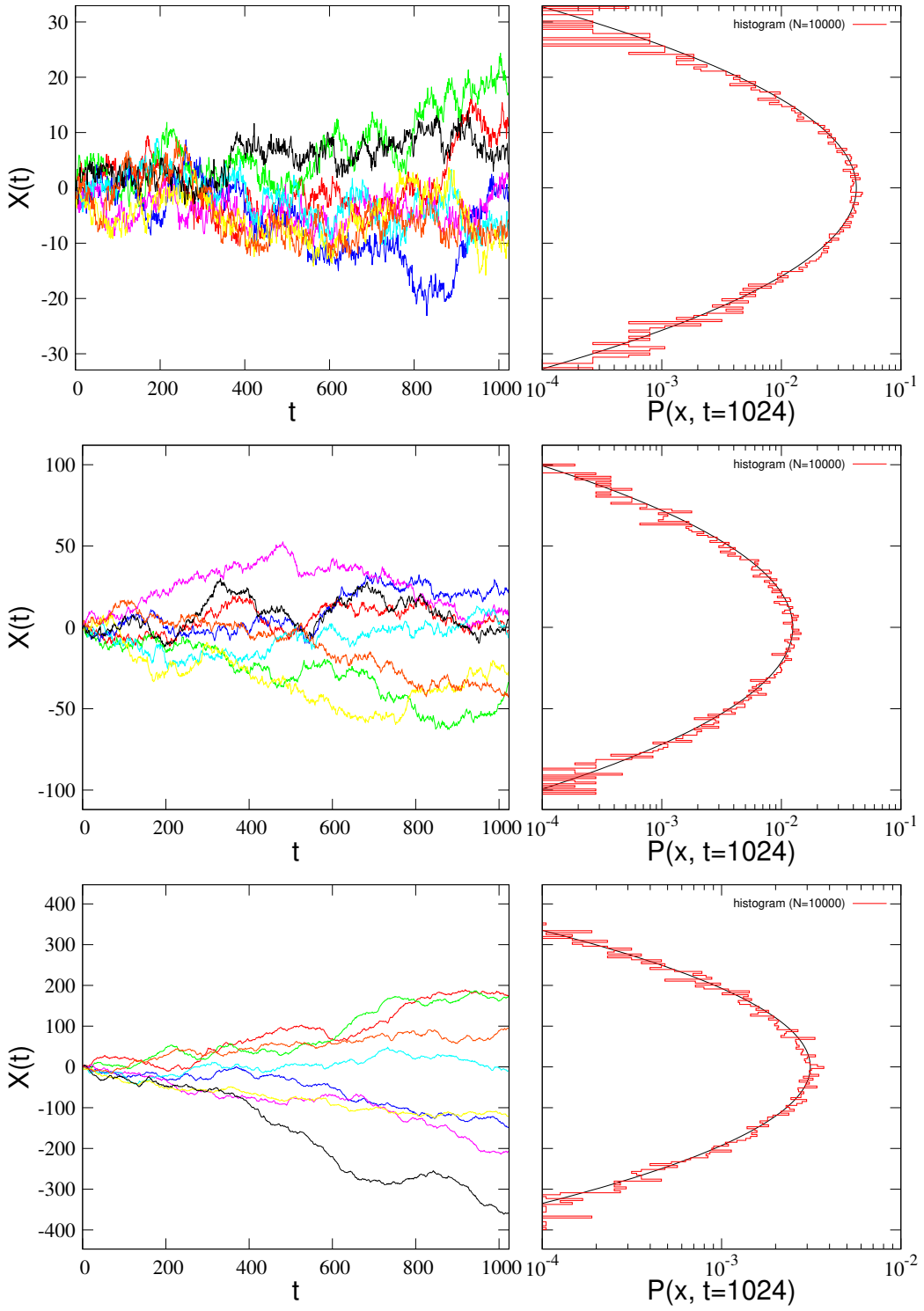


Figure 4.3: Left: Eight trajectories of fBm. Right: Computed one-time one-point pdf with 10^4 trajectories (red steps) at $t = 1024$ and theoretical value (4.38) (black line). From top to bottom: $H = 0.3$, 0.5, and 0.7.

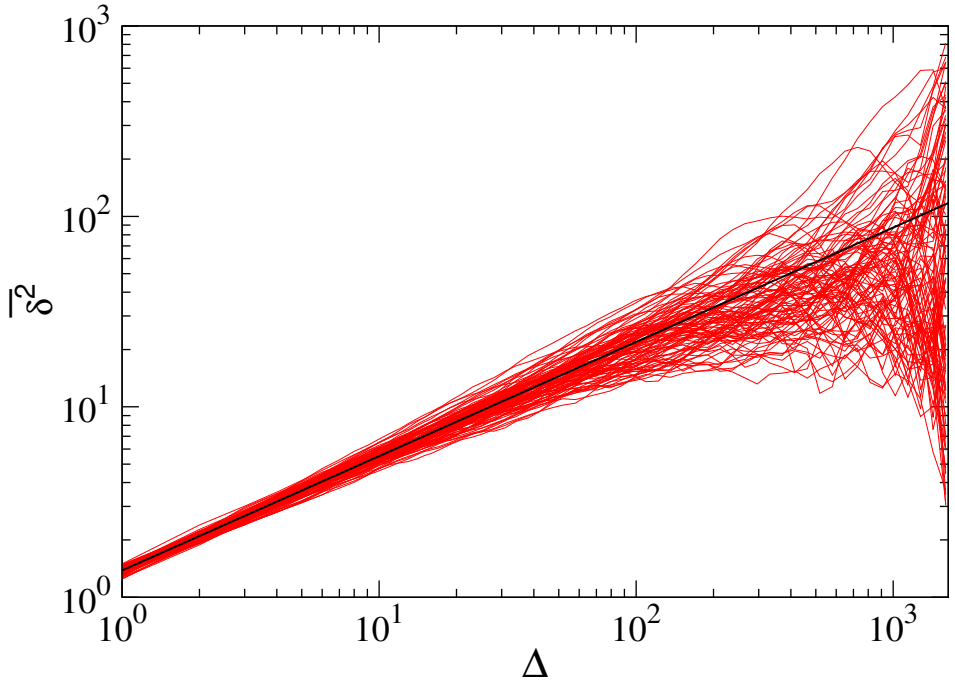


Figure 4.4: TAMSD of fBm as a function of the lag time Δ for 100 trajectories measured with $t_a = 0$ and $T = 1638$ (red lines). Black line is the theoretical EATAMSD (4.40).

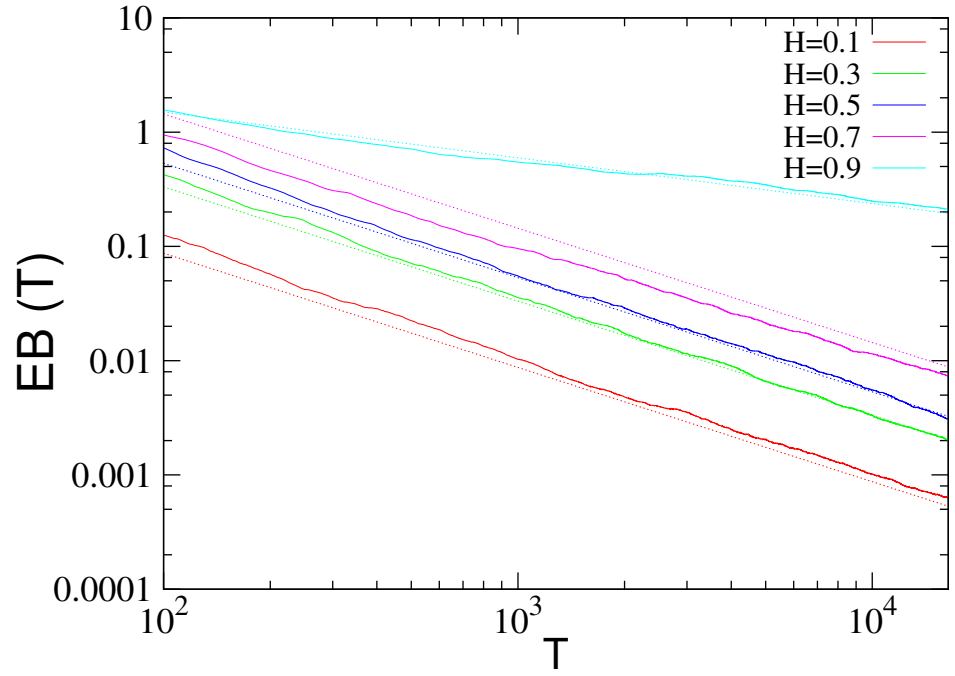


Figure 4.5: Ergodicity Breaking parameter of fBm as a function of the measurement time T for several H, $t_a = 0$, and $\Delta = 40$. Simulations (solid lines) are averaged using 10^3 trajectories. The asymptotic behavior for large T (4.42) is also shown (dashed lines).

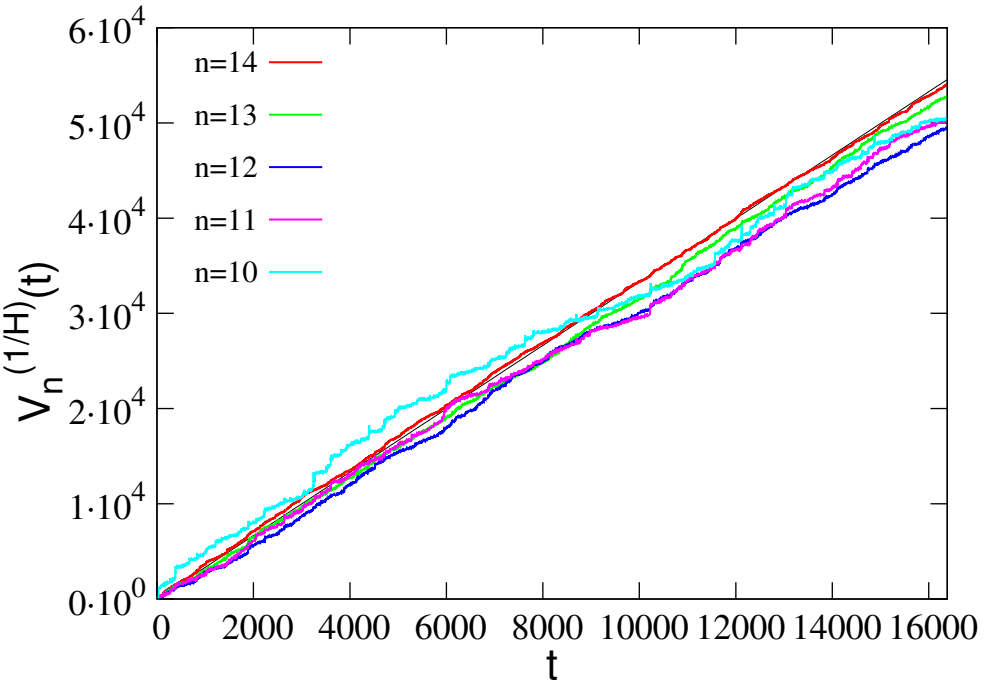


Figure 4.6: Sample p-variation of a single trajectory of fBm ($H = 0.3$) for $p = 1/H = 1/0.3$. Black line is the theoretical p-variation (4.44).

where (not equivalently written in the previous references an others) [51]

$$k = \left\langle |B_H(1)|^{1/H} \right\rangle = \frac{(2\sigma^2 V_H)^{\frac{1}{2H}}}{\sqrt{\pi}} \Gamma\left(\frac{1}{2H} + \frac{1}{2}\right). \tag{4.45}$$

It means that $V_n^{(p)}(t)$ of fBm only converges to a function for $p = 1/H$. In Figure 4.6, it can be observed how $V_n^{(1/H)}(t)$ converges to the known straight line when increasing n. On the contrary, in Figure 4.7 it is depicted how $V_n^{(2)}(t)$ diverges when increasing n. fBm can be expressed as a Langevin equation

$$\frac{dB_H(t)}{dt} = B'_H(t) \tag{4.46}$$

where $B'_H(t)$ is fractional Gaussian noise which has autocorrelation function [39]

$$\left\langle B'_H(t) B'_H(s) \right\rangle = \sigma^2(2H-1)HV_H|t-s|^{2H-2} + 2\sigma^2 HV_H|t-s|^{2H-1}\delta(t-s). \tag{4.47}$$

$B'_{1/2}(t)$ is white Gaussian noise $B'(t)$.

Aging

fBm does not present aging due to its stationary increments

$$\left\langle [B_H(t+\Delta) - B_H(t)]^2 \right\rangle = \sigma^2 V_H \Delta^{2H}. \tag{4.48}$$

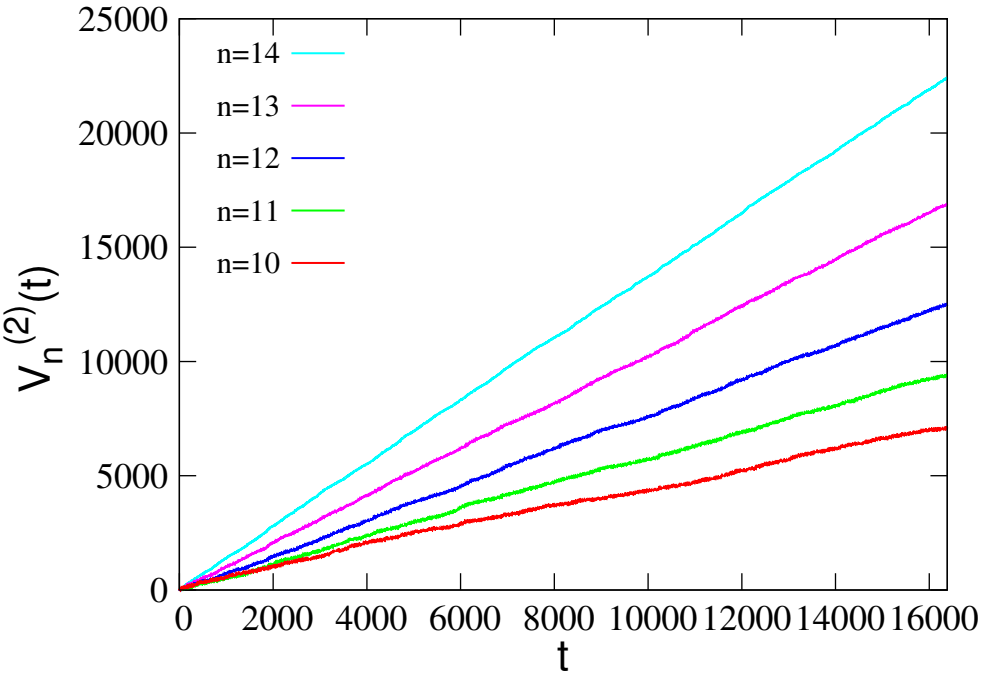


Figure 4.7: Sample p-variation of a single trajectory of fBm ($H = 0.3$) for $p = 2$.

4.3.3 Subdiffusive continuous-time random walk

The continuous-time random walk (CTRW) was originally introduced by Montroll, Weiss, and Scher. Consider a particle that has to wait a random time τ before making a jump such that the waiting times are distributed according to a distribution $\psi(\tau)$. The length of the jump δx comes also from a distribution $\lambda(\delta x)$. After the jump, a new waiting time occurs, then a new jump, and so on. Jumps and waiting times are independent of the previous values. The key factors that condition the model are the characteristic waiting time

$$\langle \tau \rangle = \int_0^\infty \tau \psi(\tau)\, d\tau \tag{4.49}$$

and the variance of the lenght of the jumps

$$\left\langle (\delta x)^2 \right\rangle = \int_{-\infty}^{\infty} (\delta x)^2 \lambda(\delta x)\, d(\delta x). \tag{4.50}$$

If they both are finite, this process corresponds to Bm $B(t)$ in the diffusion limit with $\sigma^2 = \langle (\delta x)^2 \rangle / \langle \tau \rangle$, being the details of both pdf irrelevant. The case that leads to subdiffusion is when $\langle (\delta x)^2 \rangle$ is finite but $\langle \tau \rangle$ is divergent. It is achieved with a long-tailed pdf with asymptotic behaviour [34, 85] (again, the specific form of the pdf is not important)

$$\psi(\tau) \sim \frac{\tau_0^\alpha}{|\Gamma(-\alpha)|\tau^{1+\alpha}}, \qquad 0 < \alpha < 1. \tag{4.51}$$

In Figure 4.8 different trajectories of subdiffusive CTRW are shown when the distribution of waiting times has asymptotic behaviour $\psi(\tau) \sim \alpha\tau_0^\alpha/\tau^{1+\alpha}$ for long τ. Since the rest of the figures in this section are based on simulations with the same pdf, a pre-factor correction will be used when theoretical values that involve τ_0 are shown. The EAMSD of this process $Z_\alpha(t)$ has the form

$$\left\langle Z_\alpha^2(t)\right\rangle \sim \frac{2K_\alpha}{\Gamma(1+\alpha)}t^\alpha, \tag{4.52}$$

where the generalised diffusion coefficient is

$$K_\alpha = \frac{\left\langle(\delta x)^2\right\rangle}{2\tau_0^\alpha}. \tag{4.53}$$

In Figure 4.9 the EAMSD of some simulations results is contrasted with the asymptotic expected values. In the diffusion limit, $P(x,t)$ is governed by the fractional diffusion equation

$$\frac{\partial}{\partial t}P(x,t) = {}_0D_t^{1-\alpha}K_\alpha\frac{\partial^2}{\partial x^2}P(x,t), \tag{4.54}$$

where

$${}_0D_t^{1-\alpha}P(x,t) = \frac{1}{\Gamma(\alpha)}\frac{\partial}{\partial t}\int_0^t \frac{P(x,t')}{(t-t')^{1-\alpha}}dt' \tag{4.55}$$

is the Riemann-Liouville fractional derivative operator. A closed form for $P(x,t)$ can be written in terms of the Fox function [67] or the M-Wright/Mainardi function (the equivalence can be found in reference [57])

$$P(x,t) = \frac{1}{\sqrt{4K_\alpha t^\alpha}}H_{1\,1}^{1\,0}\left[\frac{|x|}{\sqrt{K_\alpha t^\alpha}}\middle|\begin{matrix}(1-\alpha/2,\alpha/2)\\(0,1)\end{matrix}\right] \tag{4.56}$$

$$= \frac{1}{\sqrt{4K_\alpha t^\alpha}}M_{\alpha/2}\left(\frac{|x|}{\sqrt{K_\alpha t^\alpha}}\right). \tag{4.57}$$

In Figure 4.10, $P(x,t)$ is plotted. The ausence of a characteristic waiting time scale makes each trajectory no self-averaging even for very long measurement times T. It has as a consequence a disparity between EAMSD and TAMSD, known as weak ergodictiy breaking. Knowing that the mean value of the number of jumps between 0 and t behaves as [34]

$$\langle n(t)\rangle \sim \frac{t^\alpha}{\tau^\alpha\,\Gamma(1+\alpha)}, \tag{4.58}$$

the EATAMSD can be calculated and is [12, 34, 49]

$$\left\langle\overline{\delta^2(\Delta)}\right\rangle \sim \frac{2K_\alpha}{\Gamma(1+\alpha)}\frac{\Delta}{T^{1-\alpha}}. \tag{4.59}$$

As you can see, EATAMSD is very different from EAMSD except in the limit $\alpha \to 1$, indicating weak ergodicity breaking. In Figure 4.11, the TAMSD of several trajectories is displayed along the EATAMSD. It can be seen how the general slope of TAMSD is the usually

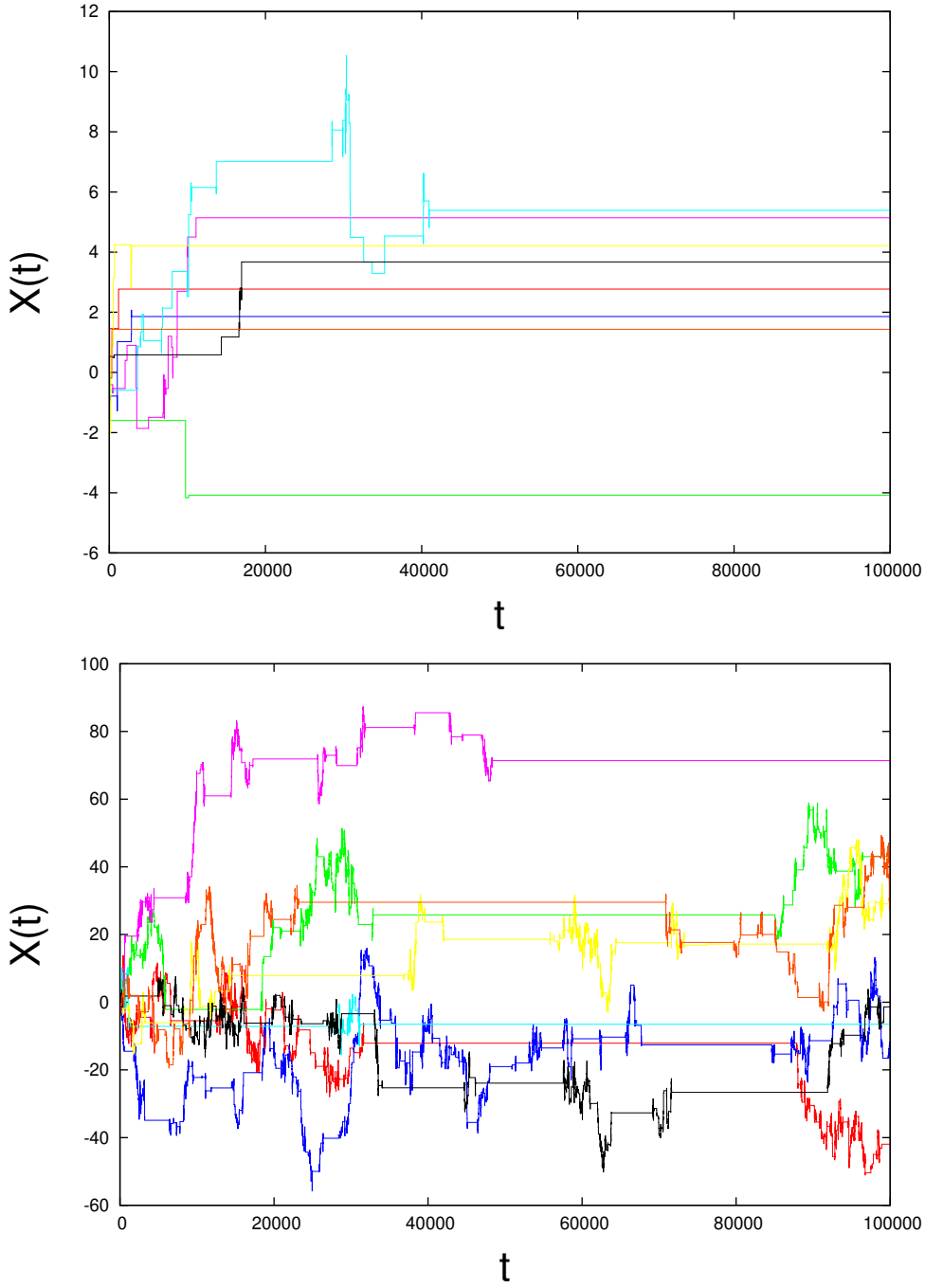


Figure 4.8: Eight trajectories of subdiffusive CTRW with a Gaussian distribution of jump sizes with $\langle(\delta x)^2\rangle = 1$ and Maneville-Pomeau distribution of waiting times $\psi(\tau) = \alpha/(\tau_0(1+\tau/\tau_0)^{1+\alpha})$ with $\tau_0 = 1$. From top to bottom: $\alpha = 0.2$ and $\alpha = 0.8$.

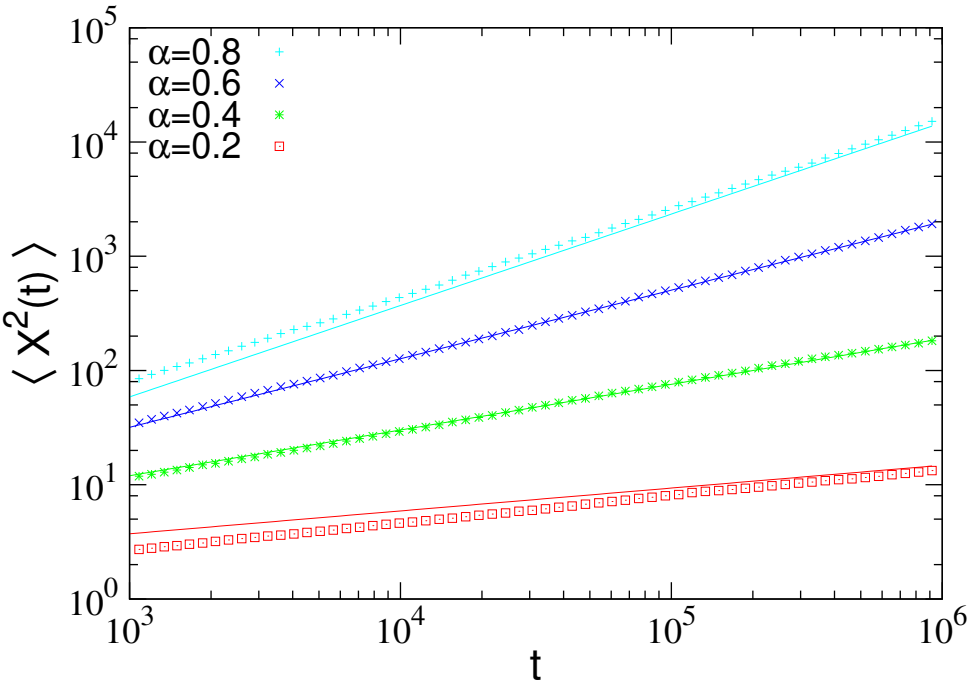


Figure 4.9: EAMSD of subdiffusive CTRW with a Gaussian distribution of jump sizes with $\langle(\delta x)^2\rangle = 1$ and Maneville-Pomeau distribution of waiting times $\psi(\tau) = \alpha/(\tau_0(1+\tau/\tau_0)^{1+\alpha})$ with $\tau_0 = 1$. Points are simulation results with $N = 10^4$, and solid lines are asymptotic theoretical values (4.52).

the same than the EATAMSD, but with desviations specially in some trajectories (an effect that increases when α decreases). Also, the amplitudes of each TAMSD is different. Both effects are due to the ausence of a characteristic time scale of waiting times. The Ergodicity Breaking parameter is [34]

$$\begin{aligned}\lim_{T\to\infty} E_B(T) &= \frac{2\Gamma^2(1+\alpha)}{\Gamma(1+2\alpha)} - 1 \\ &= \alpha\frac{\Gamma^2(\alpha)}{\Gamma(2\alpha)} - 1.\end{aligned} \tag{4.60}$$

The Ergodicity Breaking parameter has been plotted as a function of T in Figure 4.12. The subdiffusive CTRW $Z_\alpha(t)$ can be also written as a subordination process [51, 53, 54, 61]

$$Z_\alpha(t) = \sqrt{2K_\alpha}B(S_\alpha(t)) \tag{4.61}$$

where $B(t)$ is the Bm and $S_\alpha(t)$ is the inverse α-stable subordinator independent of $B(t)$ defined as

$$S_\alpha(t) = \inf\{\tau > 0 : U_\alpha(\tau) > \tau\}, \tag{4.62}$$

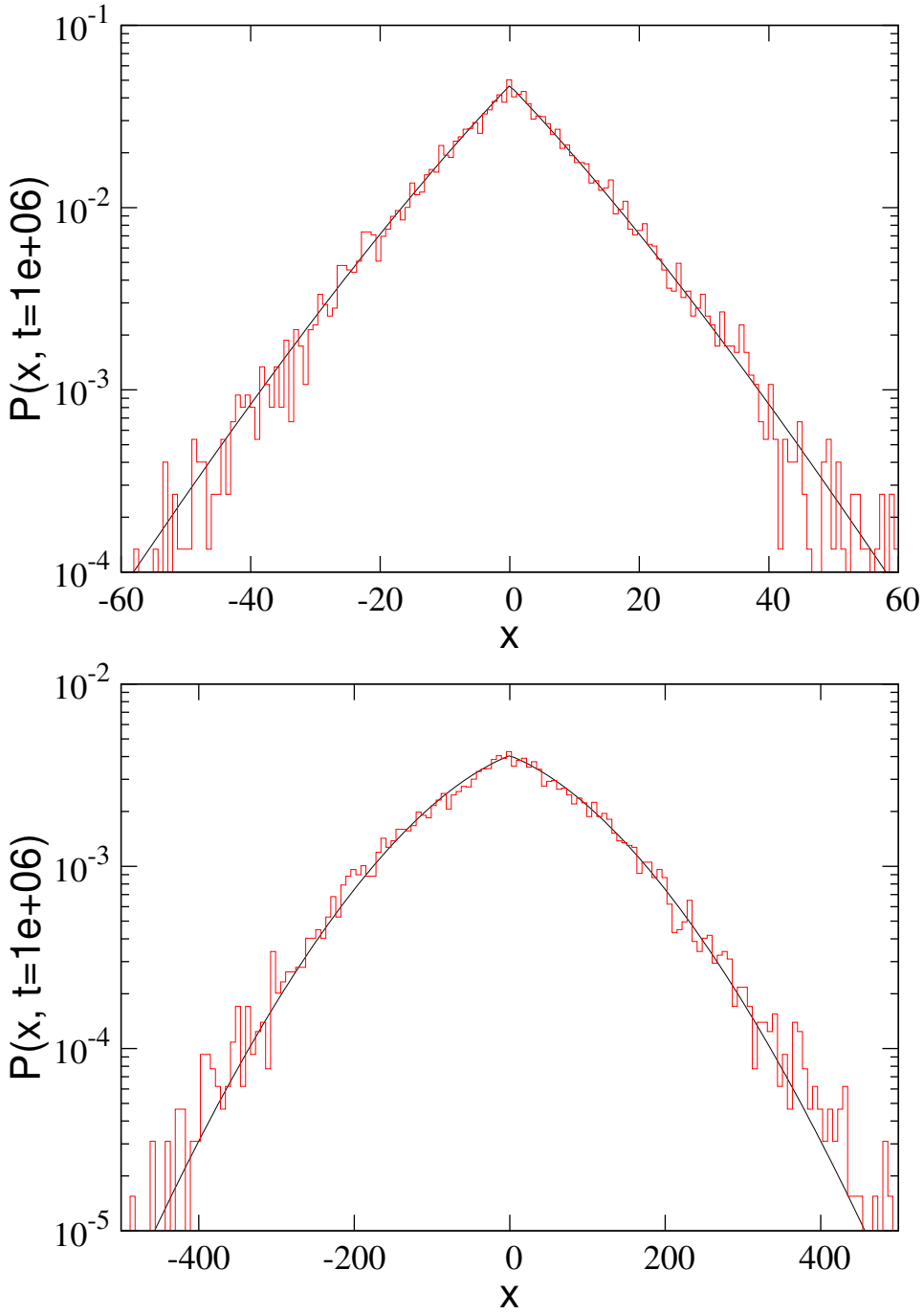


Figure 4.10: One-time one-point pdf of subdiffusive CTRW with a Gaussian distribution of jump sizes with $\langle(\delta x)^2\rangle = 1$ and Maneville-Pomeau distribution of waiting times $\psi(\tau) = \alpha/(\tau_0(1+\tau/\tau_0)^{1+\alpha})$ with $\tau_0 = 1$. Red line is the histogram for $N = 10^4$ while black line is the theoretical value (4.56). Parameters: $t = 10^6$, $\alpha = 0.4$ (top) and 0.8 (bottom).

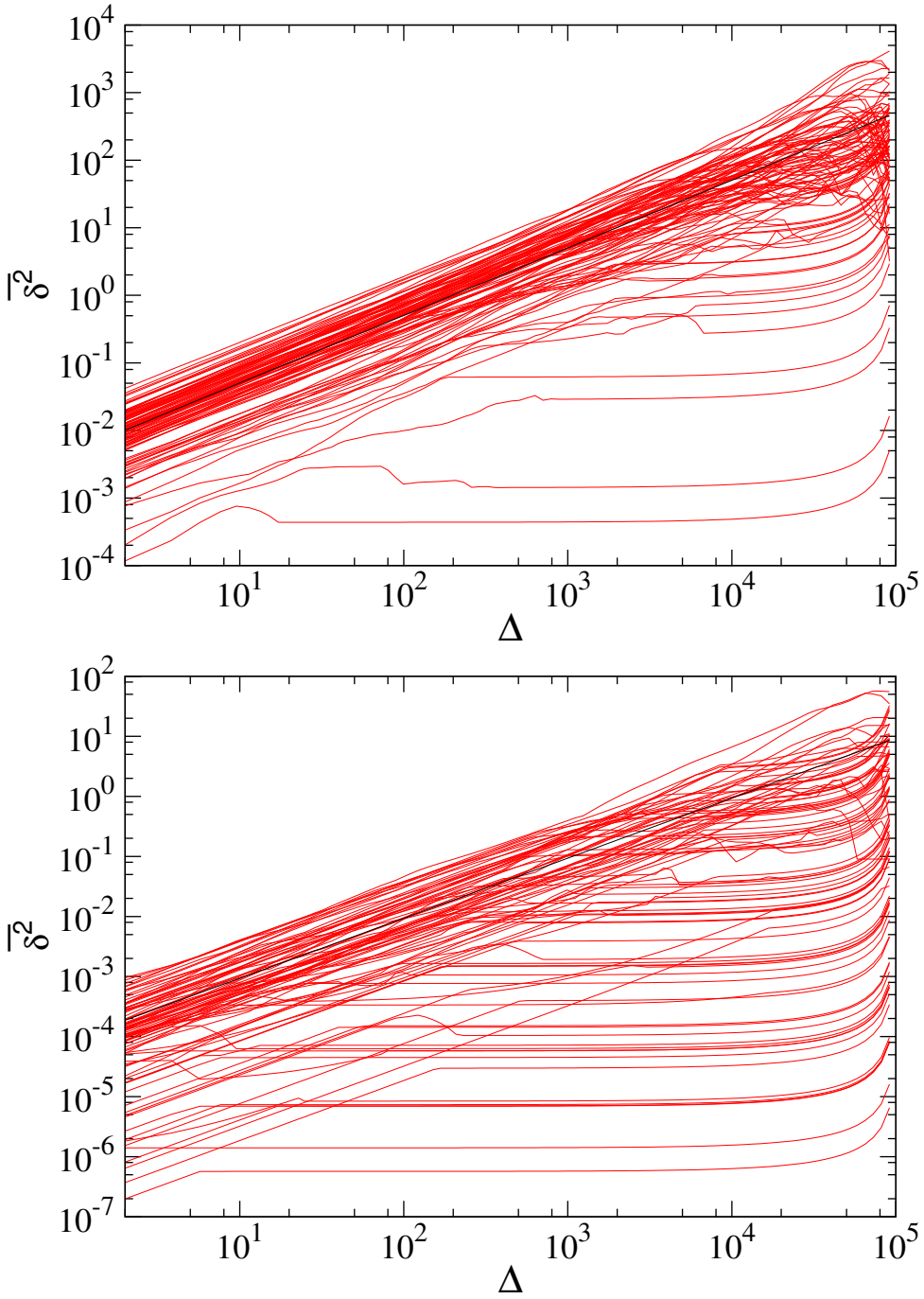


Figure 4.11: TAMSD (red lines) of 100 trajectories of subdiffusive CTRW with a Gaussian distribution of jump sizes with $\langle(\delta x)^2\rangle = 1$ and Maneville-Pomeau distribution of waiting times $\psi(\tau) = \alpha/(\tau_0(1+\tau/\tau_0)^{1+\alpha})$ with $\tau_0 = 1$. Black line is the asymptotic EATAMSD (4.59) for $T \gg \Delta$. Parameters: $t_a = 0$, $T = 99998.7$. $\alpha = 0.6$ (top) and 0.2 (bottom).

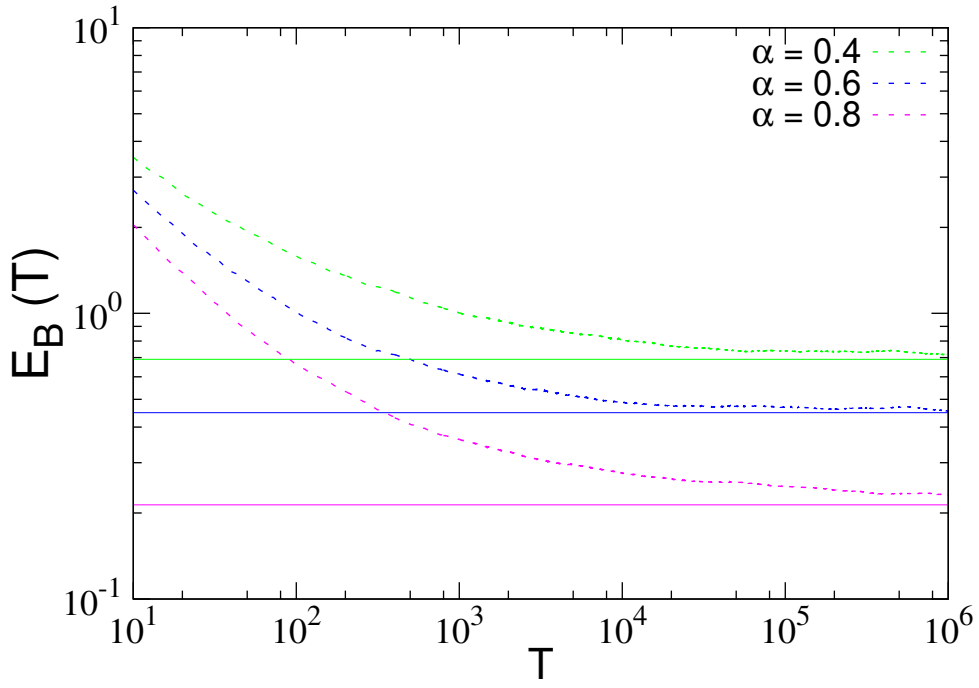


Figure 4.12: Ergodicity Breaking parameter as a function of the measurement time T of subdiffusive CTRW with a Gaussian distribution of jump sizes with $\langle(\delta x)^2\rangle = 1$ and Maneville-Pomeau distribution of waiting times $\psi(\tau) = \alpha/(\tau_0(1+\tau/\tau_0)^{1+\alpha})$ with $\tau_0 = 1$. Dashed lines show the obtained $E_B(T)$ averaged over 10^4 trajectories for $\Delta = 1.9$. Solid lines show the theoretical limit for $T \to \infty$ (4.60).

being $U_\alpha(\tau)$ a strictly increasing α-stable Lévy motion, i.e., an α-stable process which pdf fulfills

$$u(t,\tau) = \frac{1}{\tau^{1/\alpha}} u\left(\frac{t}{\tau^{1/\alpha}}\right), \qquad u(t) = u(t,1), \tag{4.63}$$

satisfying

$$\left\langle e^{-kU_\alpha(\tau)} \right\rangle = \int_0^\infty e^{-kt} u(t,\tau)\, dt = e^{-\tau k^\alpha}. \tag{4.64}$$

p-variation has the form [51, 53]

$$V^{(p)}(t) = \begin{cases} +\infty, & p < 2, \\ 2K_\alpha S_\alpha(t), & p = 2, \\ 0, & p > 2. \end{cases} \tag{4.65}$$

It means that $V_n^{(p)}(t)$ converges to a function only for $p = 2$ (see Figure 4.13). On the contrary, for $p = 1/H$ by using $2H = \alpha$, $\alpha \in (0,1)$ (doing an equivalence between the power-law exponent of the EAMSD of fBm and this CTRW), $V_n^{(p)}(t)$ converges to zero when $n \to \infty$ (see Figure 4.14).

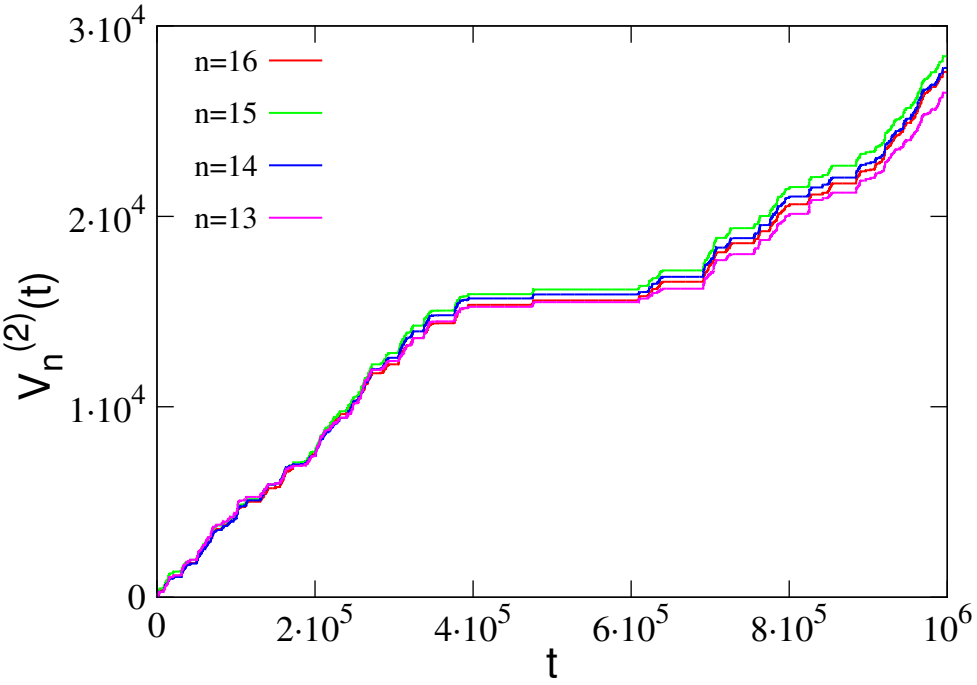


Figure 4.13: Sample p-variation of a single trajectory of subdiffusive CTRW ($\alpha = 0.8$) for $p = 2$.

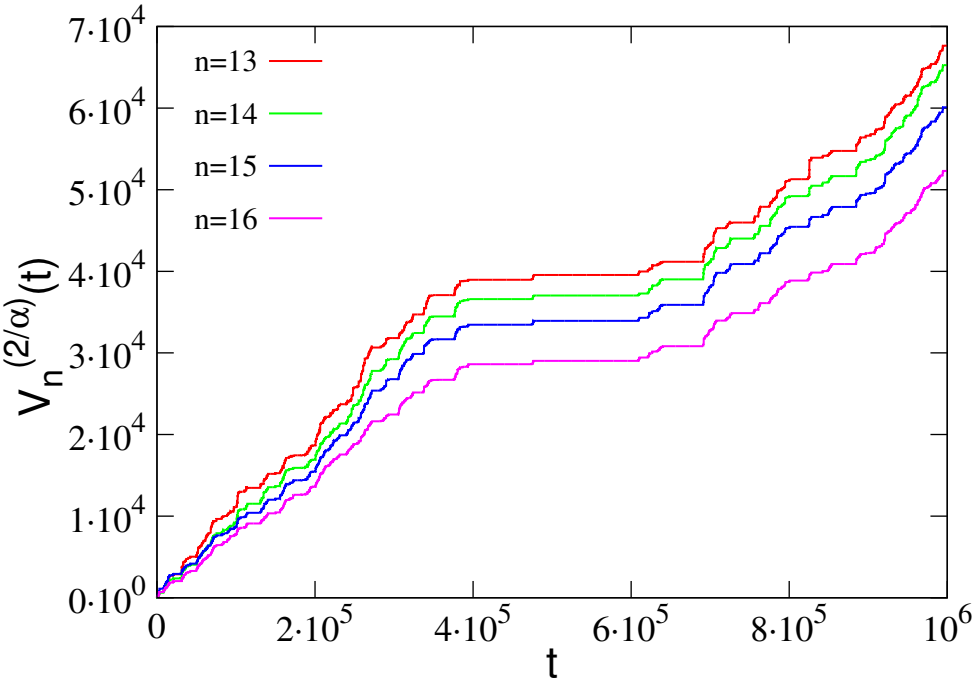


Figure 4.14: Sample p-variation of a single trajectory of subdiffusive CTRW ($\alpha = 0.8$) for $p = 2/\alpha = 2.5$.

Aging

If measures are started at $t_a > 0$, observables of CTRW differ. It is the case of the EAMSD. If $t, t_a > \tau_0$ [3],

$$\left\langle Z_\alpha^2(t)\right\rangle_a \sim \frac{2K_\alpha}{\Gamma(1+\alpha)}[(t+t_a)^\alpha - t_a^\alpha], \tag{4.66}$$

that can be approximated to

$$\left\langle Z_\alpha^2(t)\right\rangle_a \sim \begin{cases} \frac{2K_\alpha}{\Gamma(1+\alpha)} t^\alpha, & t \gg t_a, \\ \frac{2\alpha K_\alpha}{\Gamma(1+\alpha)} \frac{t}{t_a^{1-\alpha}}, & t_a \gg t. \end{cases} \tag{4.67}$$

The EATAMSD is also affected by aging [85]

$$\left\langle \overline{\delta_a^2(\Delta)} \right\rangle \sim \frac{2K_\alpha}{\Gamma(1+\alpha)} \Lambda_\alpha(t_a/T) \frac{\Delta}{T^{1-\alpha}} \tag{4.68}$$

$$= \Lambda_\alpha(t_a/T) \left\langle \overline{\delta^2(\Delta)} \right\rangle, \tag{4.69}$$

where

$$\Lambda_\alpha(z) = (1+z)^\alpha - z^\alpha. \tag{4.70}$$

In the case $t_a \gg T$, it is obtained the equivalence with $\left\langle Z_\alpha^2(\Delta)\right\rangle_a$ for $t_a \gg \Delta$

$$\left\langle \overline{\delta_a^2(\Delta)} \right\rangle \sim \frac{2\alpha K_\alpha}{\Gamma(1+\alpha)} \frac{\Delta}{t_a^{1-\alpha}}, \qquad t_a \gg T. \tag{4.71}$$

4.3.4 Fractional Langevin Equation

An alternative to Bm is the Langevin equation [46]

$$m\frac{d^2x}{dt^2} = -\gamma^* \frac{dx}{dt} + B'(t), \tag{4.72}$$

where $B'(t)$ is white Gaussian noise, γ^* is a friction coefficient, and m is the mass of test particle. When the random noise is non-white, the motion is described by the generalized Langevin equation (GLE) [39]

$$m\frac{d^2x}{dt^2} = -\gamma^* \int_0^t \mathcal{K}(t-t') \frac{dx(t')}{dt'} dt' + \xi(t), \tag{4.73}$$

where $\xi(t)$ is the noise and $\mathcal{K}$ the memory kernel, satisfying the fluctuation-dissipation theorem

$$\langle \xi(t)\xi(t')\rangle = k_B T \mathcal{K}(t-t'). \tag{4.74}$$

where T here is the temperature (not to confuse with the measurement time, that in this section we will call T_m). When the noise is fractional Gaussian noise $B'_H(t)$, $\mathcal{K}$ decays as a

power-law and the GLE becomes the fractional Langevin equation (FLE)

$$m\frac{d^2x}{dt^2} = -\gamma^* \int_0^t (t-t')^{2H-2}\frac{dx(t')}{dt'}dt' + \eta^* B'_H(t), \tag{4.75}$$

where $1/2 < H < 1$ (to avoid that the memory integral diverge) and η^* is the noise amplitude that is

$$\eta^* = \sqrt{\frac{\gamma^* k_B T}{\sigma^2 V_H H(2H-1)}}. \tag{4.76}$$

Supposing that $\left.\frac{dx(t)}{dt}\right|_{t=0} = v_0$ and $x(0) = 0$, it is found that the relaxation dynamics follows the form

$$\langle x(t)\rangle = v_0\, t\, E_{2H,2}\left(-\frac{\gamma^*\Gamma(2H-1)}{m}t^{2H}\right) \tag{4.77}$$

$$\sim \frac{v_0 m}{\gamma^*\Gamma(2H-1)\Gamma(2-2H)}t^{1-2H}, \tag{4.78}$$

being $E_{\alpha,\beta}(z)$ the generalized Mittag-Leffler function. The EAMSD is

$$\left\langle x^2(t)\right\rangle = \frac{2k_BT}{m}t^2 E_{2H,3}\left(-\frac{\gamma^*\Gamma(2H-1)}{m}t^{2H}\right) \tag{4.79}$$

$$\sim \frac{2k_BT}{\gamma^*\Gamma(2H-1)\Gamma(3-2H)}t^{2-2H}, \tag{4.80}$$

where it is supposed that $\langle v_0^2\rangle = k_BT/m$. For short times, $\langle x^2(t)\rangle \sim (k_BT/m)t^2$. The process is ergodic since the TAMSD behaves as [66]

$$\overline{\delta^2(\Delta)} = \left\langle x^2(\Delta)\right\rangle, \qquad \Delta \ll T_m, \tag{4.81}$$

being T_m the measurement time as it was said before. In reference [26] it is demonstrated how, except some changes like $2H \to 2-2H$, the Ergodicity Breaking parameter behaves similarly to (4.42) when $T_m \to \infty$.

4.3.5 Superstatistics

Another way to obtain anomalous diffusion is called superstatistics [4, 5]. It is based on the idea that the dynamics is a mixture of Gaussian processes (which probability density function is Gaussian) with different diffusivities caused by an inhomogeneous media. This way, studying a single trajectory only the properties of the Gaussian process are recovered, but when considering the ensemble the characteristics are very different. In particular, the probability density function is weighted by the distribution of local diffusivities. It also appears ergodicity breaking.

In [89], ggBm (which characterization will be the main scope of Chapter 5) is compared with superstatistics, which are apparently very similar. It is stated that the ggBm is defined through the explicit construction of a probability space based on self-similar increments and

in the case that the random variable of its representation is undestood as a diffusivity, it leads to an interpretation complementary to superstatistics.

4.4 Derivation of MSD of fBm

Let us calculate the EAMSD of fBm starting with expresion (4.33) to simplify the initial calculations. We are going to demonstrate it for $t > 0$ because it is our case of interest. Defining

$$f_t(t') := \frac{1}{\Gamma(H+1/2)}\left[(t-t')_+^{H-1/2} - (-t')_+^{H-1/2}\right], \tag{4.82}$$

we can write

$$\begin{aligned}
\left\langle B_H^2(t)\right\rangle &= \left\langle \int_{-\infty}^{\infty} f_t(t')B(t')dt' \int_{-\infty}^{\infty} f_t(t'')B(t'')dt'' \right\rangle && (4.83)\\
&= \left\langle \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_t(t')f_t(t'')B(t')B(t'')dt'dt'' \right\rangle && (4.84)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_t(t')f_t(t'')\left\langle B(t')B(t'')\right\rangle dt'dt'' && (4.85)\\
&= \sigma^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_t(t')f_t(t'')\delta(t'-t'')dt'dt'' && (4.86)\\
&= \sigma^2\int_{-\infty}^{\infty} f_t^2(t')dt' && (4.87)\\
&= \frac{\sigma^2}{\Gamma^2(H+1/2)}\int_{-\infty}^{\infty}\left[(t-t')_+^{H-1/2} - (-t')_+^{H-1/2}\right]^2 dt' && (4.88)
\end{aligned}$$

Now, by applying definition of operator $(\cdot)_+$ we can separate the integral in the following way

$$\left\langle B_H^2(t)\right\rangle = \frac{\sigma^2}{\Gamma^2(H+1/2)}\left[\int_{-\infty}^{0}\left[(t-t')^{H-1/2} - (-t')^{H-1/2}\right]^2 dt' + \int_0^t (t-t')^{2H-1}dt'\right] \tag{4.89}$$

Applying the change of variables $x = -t'/t$ to the first integral and $x = (t-t')/t$ to the second one, it results

$$\begin{aligned}
\left\langle B_H^2(t)\right\rangle &= \frac{\sigma^2 t^{2H}}{\Gamma^2(H+1/2)}\left[\int_{-\infty}^{0}\left[(1+x)^{H-1/2} - x^{H-1/2}\right]^2 dx + \int_0^1 x^{2H-1}dx\right] && (4.90)\\
&= \sigma^2 V_H t^{2H} && (4.91)
\end{aligned}$$

where

$$V_H := \frac{1}{\Gamma^2(H+1/2)}\left[\int_{-\infty}^{0}\left[(1+x)^{H-1/2} - x^{H-1/2}\right]^2 dx + \int_0^1 x^{2H-1}dx\right]. \tag{4.92}$$

Both integrals have an explicit expression and after some manipulation using properties of the gamma function it is arrived to (4.36).

4.5 Simulation of ggBm

To simulate the fBm we followed the exact Hosking method described in [27]. This method (also known as Durbin or Levinson) simulates a general stationary Gaussian process, in this case fractional Gaussian noise, and then takes cumulatives sums to obtain fractional Brownian motion trajectories.

For the obtention of ggBm, it is necessary to multiply each trajectory of a fractional Brownian motion process by the square root of a constant extracted from a random variable Λ_β distributed according to the one-side M-Wright/Mainardi function (see Chapter 5 for more details),

$$X_{\beta,H}(t) = \sqrt{\Lambda_\beta}\, B_H(t). \tag{4.93}$$

This random variable can be computed in terms of the extremal stable random variable of order β, $\mathcal{L}_\beta^{\text{ext}}$ [73]

$$\Lambda_\beta = [\mathcal{L}_\beta^{\text{ext}}]^{-\beta}, \tag{4.94}$$

that can be generated by the method of Chambers, Mallows and Stuck [15]

$$\mathcal{L}_\beta^{\text{ext}} = \frac{\sin[\beta(r_1+\pi/2)]}{(\cos r_1)^{1/\beta}} \left(\frac{\cos[r_1 - \beta(r_1+\pi/2)]}{-\ln r_2} \right)^{(1-\beta)/\beta}, \qquad 0<\beta<1, \tag{4.95}$$

where r_1 and r_2 are uniformly distributed random variables in $(-\pi/2, \pi/2)$ and $(0,1)$, respectively.

4.6 Community ecology

4.6.1 Usual neutral models

Voter Model with speciation

The voter model [48] has been widely studied, from spreading of epidemic diseases [76] to opinion dynamics [13], linguistics [23] and spatial conflicts [20]. The voter model with speciation [29] (or multi-species voter model) is a generalization of the voter model. It is defined on a lattice where each node has one individual that belongs to a specific species. At each time step, one individual of the lattice is removed, what is known as a death event. Then, a new individual replaces the removed one. The species associated to this new individual is chosen according to the following rule: With probability ν, it is a new species not present in the system. That is known as a speciation event. With probability $1-\nu$, it becomes the species of a neighbour, what is known as a reproduction event. The procedure is illustrated in Figure 4.15. The parent individual is selected according to a dispersal kernel $P(\vec{r})$ (a probability distribution). Typically, the dispersal kernel is such that one the nearest-neighbours (NN) is chosen with the same probability.

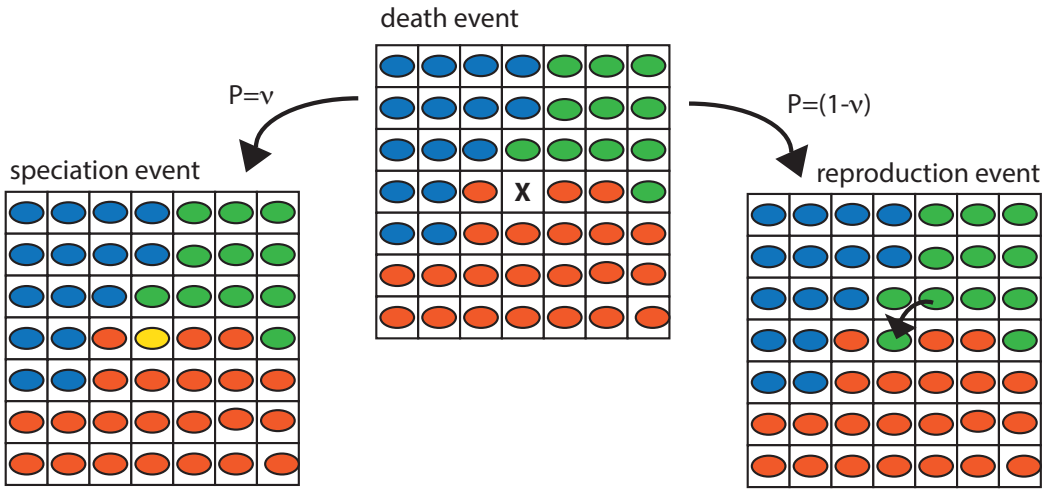


Figure 4.15: Different cases after a death event in voter model with speciation.

Stepping-Stone Model with speciation

This model [14, 43] generalizes the voter model with speciation by considering that each node of the lattice has M individuals. At each step, an individual is randomly removed. With probability ν, it is replaced by an individual of a new species not present in the system. With probability $1-\nu$, a reproduction event happens. The parent of the individual is, with probability $1-\mu$, one of the remaining individuals in the same node, and, with probability μ, one individual belonging to a neighbour node (according to a dispersal kernel).

Contact Process with speciation

The contact process with speciation is the multi-species variant of the contact process [28, 32, 36, 37, 59]. Let us consider a lattice of individuals and a continuous time variable so each individual dies at rate d and reproduces at rate b. In this model there are two key points. Firstly, when a death event occurs, the associated node is left empty. Secondly, a reproduction event is successful only if there is at least one vacant neighbour node. In the last case, one of the empty nodes is chosen randomly. The new individual will have associated a new species with probability ν or will belong to the same species than the parent with probability $1-\nu$.

4.6.2 Ecological patterns

β-diversity

β-diversity measures how species composition varies with distance. It is the probability $F(\vec{r})$ that two random individuals separated by $\vec{r}$ belong to the same species. It can be written in terms of the two-point correlation function

$$G_{s_i,s_j}(\vec{r}) = \langle n_{s_i}(\vec{x}) n_{s_j}(\vec{x}+\vec{r}) \rangle, \tag{4.96}$$

being $n_{s_i}(\vec{x})$ the number of individuals of species s_i at position $\vec{x}$. It holds

$$F(\vec{r}) = \frac{\sum_i G_{s_i,s_i}(\vec{r})}{\sum_{i,j} G_{s_i,s_j}(\vec{r})} \tag{4.97}$$

where the sum includes all the species in the ecosystem.

Species-Area Relationship

The Species-Area Relationship (SAR) is defined as the average number of species of a given taxonomic level occupying a given area. It quantifies how habitats of different sizes support different number of species and is considered a measure of spatial biodiversity.

Species-Abundance Distributions

Species Abundance distributions (SADs) measure the relative abundance (number of individuals) of species in a given area. Let n_i be the abundance of the species i present in area A. SAD is the distribution $P(n; A)$ such that $P(n; A)dn$ represents the probability that a randomly picked species has a abundance between n and $n + dn$.

Species persistence-times

Survival or persistence time within a geographic region is the time occurring between the appearance of a given species and its local extintion.

4.7 Methodology

For the first article, the methodology has been:

- Study of the main bibliography related with ggBm.
- Develop the main software for computation of the ggBm.
- Study of the main bibliography that deals with anomalous diffusion experiments.
- Calculate the observables of interest in the previous experiments for the ggBm. It will require to:
  - Develop the needed software tools to compute the observables.
  - Calculate analytically (when possible) the quantities, which will be checked with the numerical output.
- Determine if ggBm can explain the empirical data.

For the second article, the mathematical results have been obtained and checked by:

- Several repetitions of the calculations.
- Check known results by simplifying the findings to particular cases.
- Use of software tools to check the properties of the obtained functions.

For the third article, the methodology was:

- Research the state of the art in spatial neutral theory and near-neutral models.
- Extensive numerical computations of the models, specially for the voter model with speciation.
- Repeat calculations of important observables and present them in a reader-friendly way.

Chapter 5

Results Part 1: Generalised gray Brownian motion

> One cannot fight what one does not know.
>
> Rollo May, The Meaning of Anxiety

In this chapter, we will mainly deal with a particular process, the ggBm, that satisfies the conditions established in the Hypothesis section 3.1. It will be only in the Ergodicity section 5.3 where it will be obtained a property that is common for all the processes satisfying the specified conditions in 3.1.

5.1 Definition

Let us consider in particular the process known as generalised grey Brownian motion [69–71] (ggBm)

$$X_{\beta,H}(t) = \sqrt{\Lambda_\beta} B_H(t), \tag{5.1}$$

where the random variable Λ_β is positive and distributed according to the one-side M-Wright/Mainardi function [55, 72]

$$M_\beta(\lambda) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\lambda^k}{\Gamma[-\beta k + (1-\beta)]}, \tag{5.2}$$

with $\lambda \geq 0$ and $0 < \beta < 1$. B_H is the fBm described in 4.3.2. This process include as particular cases the Brownian motion when $H = 1/2$ and $\beta \to 1$, the fractional Brownian motion when $\beta \to 1$ and the grey Brownian motion when $\beta = 2H$.

5.2 Main properties

The probability density function is [71]

$$\mathcal{P}(\mathbf{x}; \gamma_H) = \frac{1}{\sqrt{(2\pi\lambda)^n} \det \gamma_H} \int_0^\infty \exp\left\{ -\frac{1}{2\lambda} \mathbf{x}^T \gamma_H^{-1} \mathbf{x} \right\} M_\beta(\lambda) d\lambda, \tag{5.3}$$

where $\mathbf{x} = (x_1, \cdots, x_n)$ and $\gamma_H = \gamma_H(t_i, t_j), i, j = 1, \cdots, n$, is the covariance matrix of the fBm

$$\gamma_H(t_i, t_j) = \frac{\sigma^2 V_H}{2}\left(t_i^{2H} + t_j^{2H} - |t_i - t_j|^{2H}\right). \tag{5.4}$$

The one-point one-time density function is

$$\begin{aligned}\mathcal{P}(x;t) &= \frac{1}{\sqrt{2\pi\lambda\sigma^2 V_H t^{2H}}}\int_0^\infty \exp\left\{-\frac{x^2}{2\lambda\sigma^2 V_H t^{2H}}\right\} M_\beta(\lambda)d\lambda \\ &= \frac{1}{\sqrt{2\sigma^2 V_H}t^H} M_{\beta/2}\left(\sqrt{\frac{2}{\sigma^2 V_H}}\frac{|x|}{t^H}\right).\end{aligned} \tag{5.5}$$

This can be observed in Figure 5.1 and 5.2 along some individual trajectories. $\mathcal{P}(x;t)$ can be written in terms of the H-function [56, 57]

$$\mathcal{P}(x;t) = \frac{1}{\sqrt{2\sigma^2 V_H}t^H} H_{1\,1}^{1\,0}\left[\sqrt{\frac{2}{\sigma^2 V_H}}\frac{|x|}{t^H}\,\middle|\begin{matrix}(1-\beta/2, \beta/2)\\ (0,1)\end{matrix}\right]. \tag{5.6}$$

The covariance matrix of the ggBm is [70, 71]

$$\gamma_{\beta,H}(t,s) = \frac{\sigma^2 V_H}{2\Gamma(1+\beta)}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \tag{5.7}$$

and the variance or EAMSD reads

$$\langle X^2_{\beta,H}(t)\rangle = \frac{\sigma^2 V_H}{\Gamma(1+\beta)}t^{2H}. \tag{5.8}$$

In addition, it satisfies

$$\left\langle [B_H(t+\Delta) - B_H(t)]^2\right\rangle = \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\Delta^{2H}. \tag{5.9}$$

The time-average MSD is

$$\overline{\delta^2(\Delta)} = \Lambda_\beta \sigma^2 V_H \Delta^{2H}, \quad T \gg \Delta, \tag{5.10}$$

and the EATAMSD

$$\left\langle \overline{\delta^2(\Delta)}\right\rangle = \frac{\sigma^2 V_H \Delta^{2H}}{\Gamma(1+\beta)}. \tag{5.11}$$

In Figure 5.3 it is shown the TAMSD of several simulated particles and the EATAMSD. Note that all of them display the same power-law with the exception of a different prefactor (different height in a log-log graph) due to the presence of the random variable Λ_β.

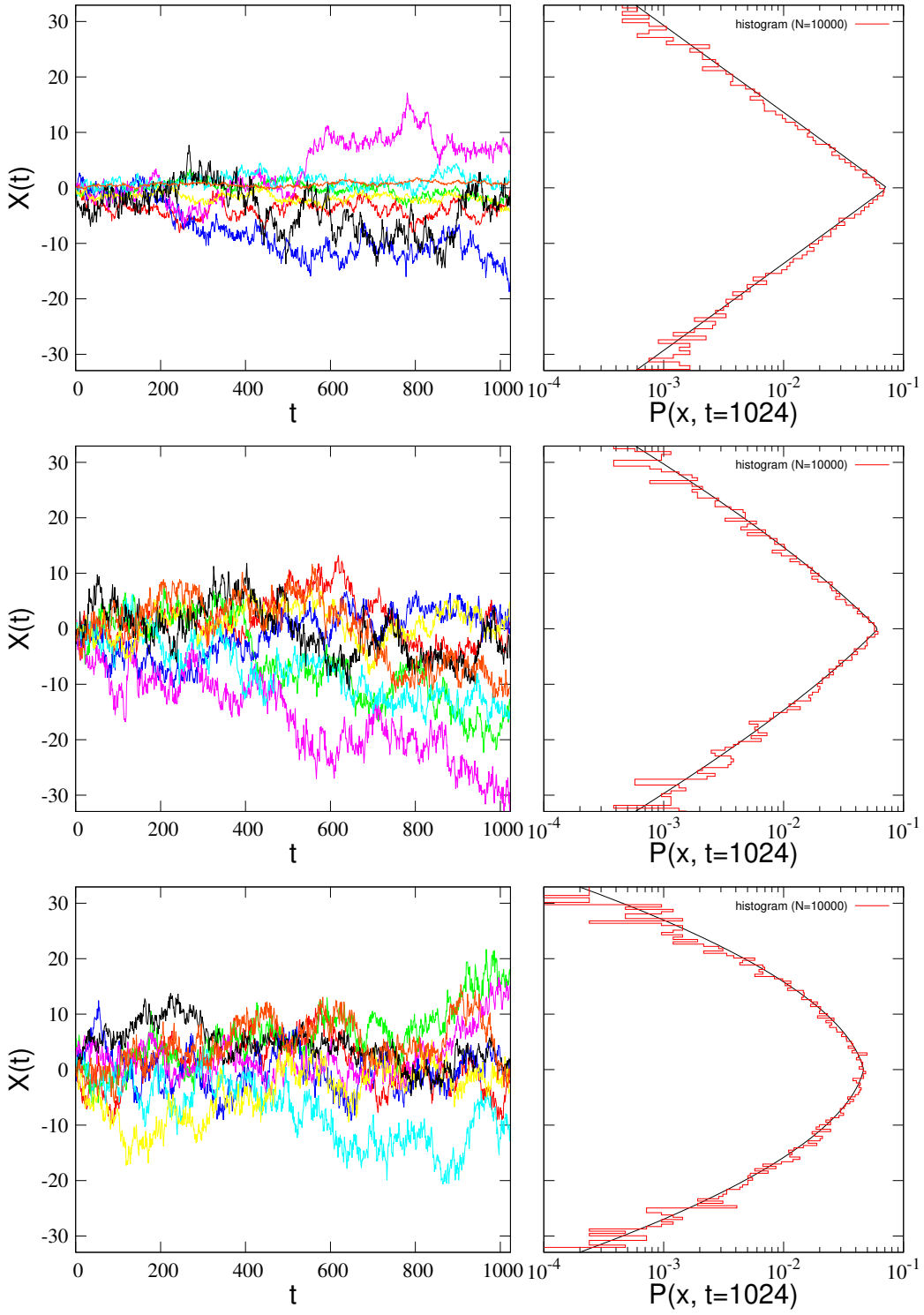


Figure 5.1: Left: eight individual trajectories of ggBm for $H = 0.3$. Right: The red line is the histogram of the one-time one-point pdf of 10^4 simulated trajectories at $t = 1024$. Black line corresponds to the theoretical curve (5.5). From top to bottom: $\beta = 0.1$, 0.5, and 0.9.

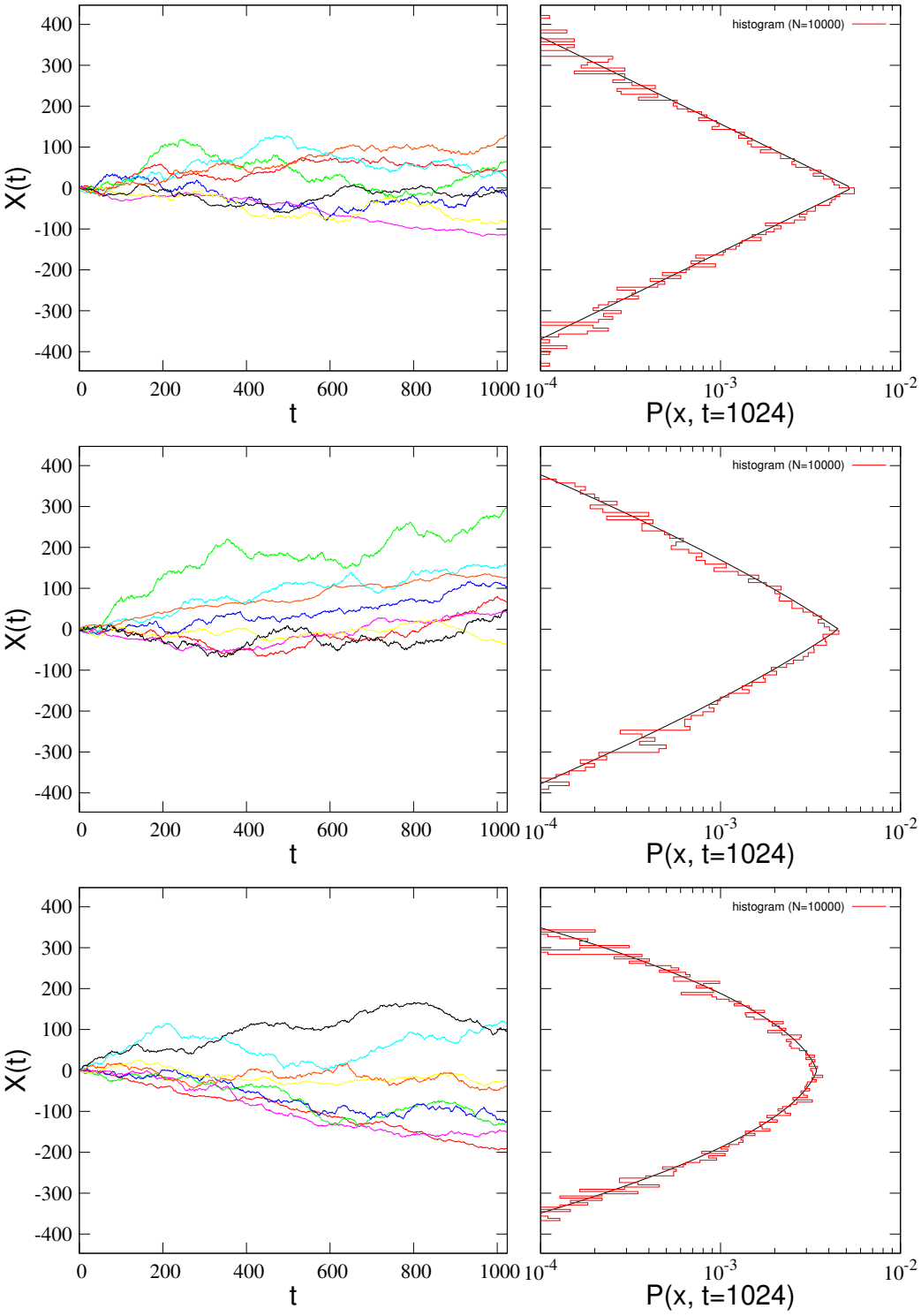


Figure 5.2: Left: eight individual trajectories of ggBm for $H = 0.7$. Right: The red line is the histogram of the one-time one-point pdf of 10^4 simulated trajectories at $t = 1024$. Black line corresponds to the theoretical curve (5.5). From top to bottom: $\beta = 0.1$, 0.5, and 0.9.

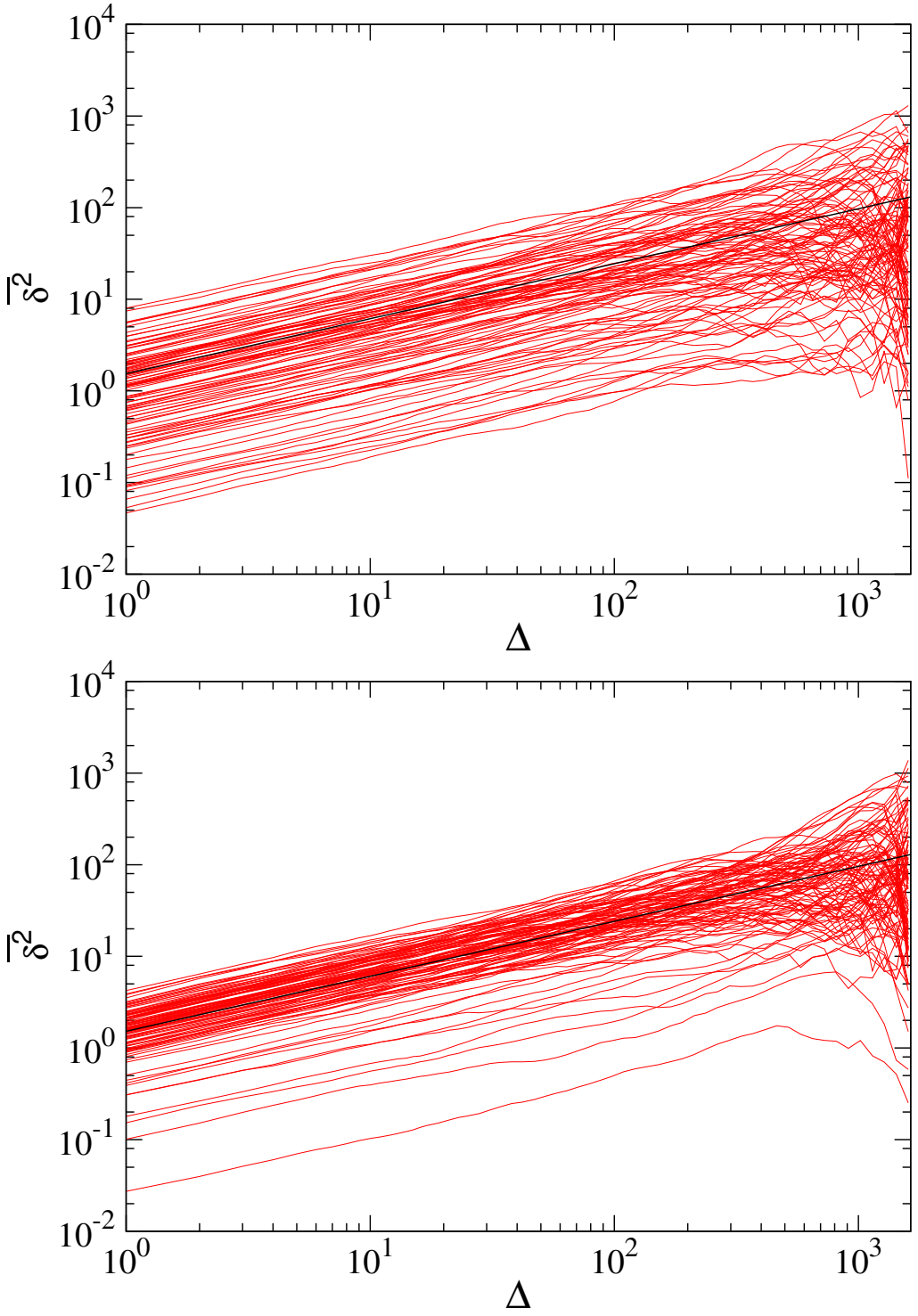


Figure 5.3: TAMSD as a function of the lag time Δ for 100 single trajectories (red lines) and EATAMSD (black line). Common parameters are $H = 0.3$, $T = 1638$ and $t_a = 0$. From top to bottom: $\beta = 0.3$ and 0.7.

5.3 Ergodicity

According to the general stochastic model considered in the Hypothesis section 3.1,

$$X(t) = l_\beta X_{\text{gen}}(t), \tag{5.12}$$

it is arrived to the conclusion that the Ergodicity Breaking parameter [26, 34]

$$E_B(T) = \frac{\left\langle \left[\overline{\delta^2}(T)\right]^2 \right\rangle}{\left\langle \overline{\delta^2}(T) \right\rangle^2} - 1 \tag{5.13}$$

tends to

$$\lim_{T\to\infty} E_B(T) = \frac{\langle l_\beta^4 \rangle}{\langle l_\beta^2 \rangle^2} - 1. \tag{5.14}$$

The condition $\langle l_\beta^4 \rangle > \langle l_\beta^2 \rangle^2$ is met for any distribution due to the inequality $K \geq S^2 + 1$ [41], where K is the kurtosis and S the skewness. In particular, for the ggBm, it is

$$\lim_{T\to\infty} E_B(T) = \beta \frac{\Gamma^2(\beta)}{\Gamma(2\beta)} - 1. \tag{5.15}$$

Function (5.15) can be observed in Figure 5.4. Figure 5.5 shows how $E_B(T)$ converges to a constant for large T that is common for different Δ.

5.4 p-variation

p-variation is defined as [53]

$$V^{(p)}(t) = \lim_{n\to\infty} V_n^{(p)}(t) \tag{5.16}$$

where $t \in [0, T]$ and

$$V_n^{(p)}(t) = \sum_{j=0}^{2^n-1} \left| X\left(\frac{(j+1)T}{2^n} \wedge t\right) - X\left(\frac{jT}{2^n} \wedge t\right) \right|^p, \tag{5.17}$$

with $a \wedge b = \min\{a, b\}$. For the ggBm, we have

$$V_{\text{ggBm}}^{(p)}(t) = \Lambda_\beta^{p/2} V_{\text{fBm}}^{(p)}(t). \tag{5.18}$$

$V_n^{(p)}(t)$ of ggBm (as well as fBm) diverges as increasing n for $p = 2 < 1/H \approx 3.33$ and converges to a straight line for $p = 1/H$. Graphics are not shown because they are essentially the same than Figure 4.6 and 4.7. For the CTRW, the value of convergence is $p = 2$.

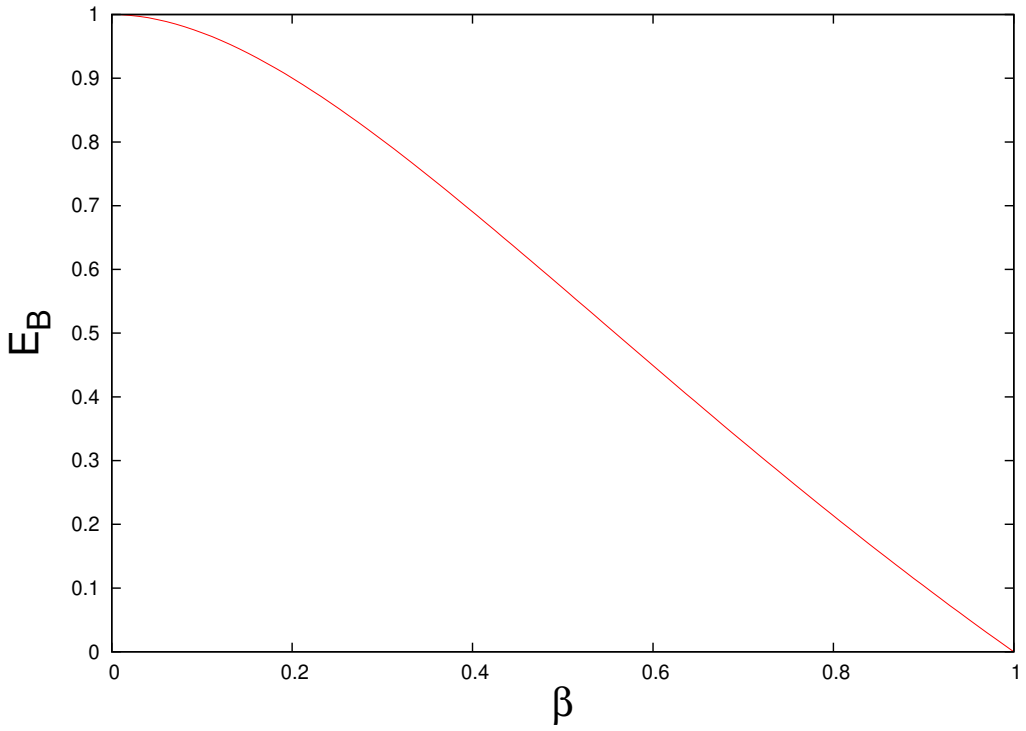


Figure 5.4: Ergodicity Breaking parameter of ggBm in the limit $T \to \infty$ (5.15) as a function of β.

5.5 Relation with experiments

The Ergodicity Breaking parameter and the p-variation are observables commonly used to characterize a stochastic process given some experimental data. There is a famous experiment describing the motion of mRNA molecules inside living E. Coli cells presented by Goldin and Cox [31]. Some features of this experiment is that it has Ergodicity Breaking, which requires models like CTRW, while presenting a p-variation fBm-like [53] (different than 0 and ∞ for $p = 1/H$), which is an eregodic stochastic process. This is the case of the ggBm, which matches both observables. More discussion on this point will be done in Chapter 8.

5.6 Aging

Sometimes, an interesting feature of a stochastic process is aging, the dependence of statistical quantitites on the initiation of the measurement time [11]. We can easily introduce aging by using a time-dependent prefactor in the formulation of our model

$$X_{\alpha,\beta,H}(t) = \sqrt{t^{\alpha}\Lambda_{\beta}}B_H(t), \qquad \alpha > -1, \tag{5.19}$$

so the covariance matrix is

$$\gamma_{\alpha,\beta,H}(t,s) = \frac{\sigma^2 V_H}{2\Gamma(1+\beta)} t^{\alpha/2}s^{\alpha/2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \tag{5.20}$$

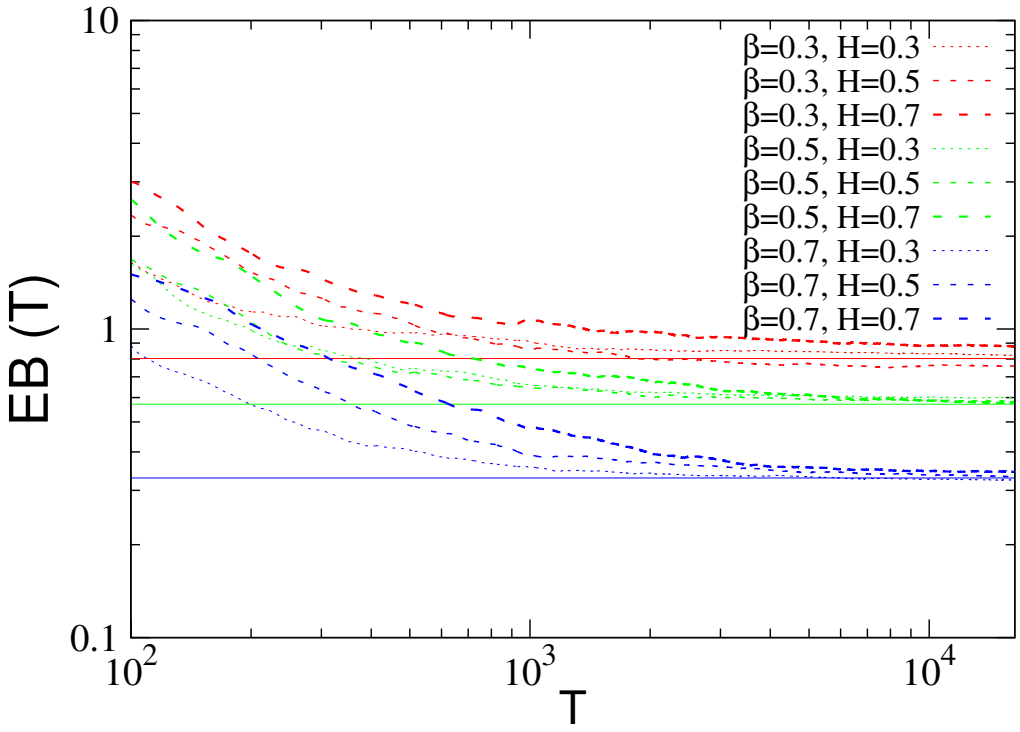


Figure 5.5: Ergodicity Breaking parameter of ggBm as a function of measurement time T, averaged over 10^3 trajectories, for $\Delta = 40$ and several values of H and β. The theoretical value of the EB (5.15) is represented with solid lines.

and the EAMSD

$$\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle = \frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{\alpha+2H}. \tag{5.21}$$

It can be shown (see section 5.7) that

$$\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a \sim \begin{cases} \frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{\alpha+2H}, & t \gg t_a, \\ \frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{2H} t_a^{\alpha}, & t_a \gg t, \end{cases} \tag{5.22}$$

and the EATAMSD is (not exact in the paper)

$$\left\langle \overline{\delta^2_a(\Delta)} \right\rangle \sim \begin{cases} \frac{\sigma^2 V_H}{(1+\alpha)\Gamma(1+\beta)} \Delta^{2H} T^{\alpha}, & T \gg \Delta, t_a, \\ \frac{\sigma^2 V_H}{\Gamma(1+\beta)} \Delta^{2H} t_a^{\alpha}, & t_a \gg T \gg \Delta. \end{cases} \tag{5.23}$$

It can be seen that $\langle X^2_{\alpha,\beta,H}(\Delta)\rangle_a = \langle \overline{\delta^2_a(\Delta)}\rangle$ when $t_a \gg T \gg \Delta$. In Figures 5.6 and 5.7, both behaviors of the EATAMSD (5.23) can be observed for different values of α.

In Table 5.1, main characteristics of fBm, subdiffusive CTRW and ggBm are compared.

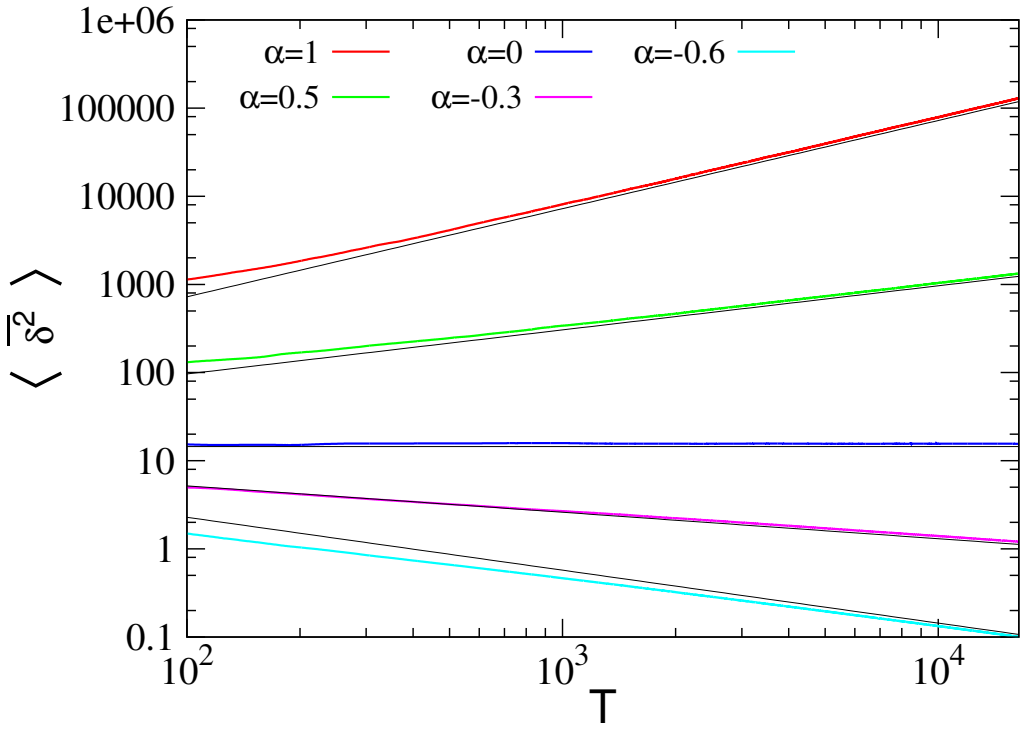


Figure 5.6: EATAMSD of the extended ggBm as a function of the measurement time T. Black lines represent the theoretical asymptotic behaviour (5.23) for $T \gg \Delta, t_a$. Parameters: $\beta = H = 0.3$, $\Delta = 50$, $t_a = 0$, $N = 10^3$.

	fBm, $H \in (0,1)$	Subdiff. CTRW, $\alpha \in (0,1)$	ggBm, $H \in (0,1)$, $\beta \in (0,1)$
P(x,t)	$\frac{e^{-\frac{x^2}{2\sigma^2 V_H t^{2H}}}}{\sqrt{2\pi\sigma^2 V_H t^{2H}}}$	$\frac{1}{\sqrt{4K_\alpha t^\alpha}} M_{\alpha/2}\left(\frac{1}{\sqrt{K_\alpha t^\alpha}}\lvert x\rvert\right)$	$\frac{1}{\sqrt{2\sigma^{2H} t^{2H}}} M_{\beta/2}\left(\sqrt{\frac{2}{\sigma^2 V_H t^{2H}}}\lvert x\rvert\right)$
$\langle X^2(t)\rangle$	$\sigma^2 V_H t^{2H}$	$\frac{2K_\alpha}{\Gamma(1+\alpha)} t^\alpha$	$\frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{2H}$
$\left\langle \overline{\delta^2(\Delta)} \right\rangle$	$\sigma^2 V_H \Delta^{2H}$	$\frac{2K_\alpha}{\Gamma(1+\alpha)T^{1-\alpha}} \Delta$	$\frac{\sigma^2 V_H}{\Gamma(1+\beta)} \Delta^{2H}$
Ergodic	Yes	No	No
p_* of p-variation	$1/H$	2	$1/H$
Aging	No	Yes	No

Table 5.1: Comparison of the main characteristics of fBm, subdiffusive CTRW and ggBm.

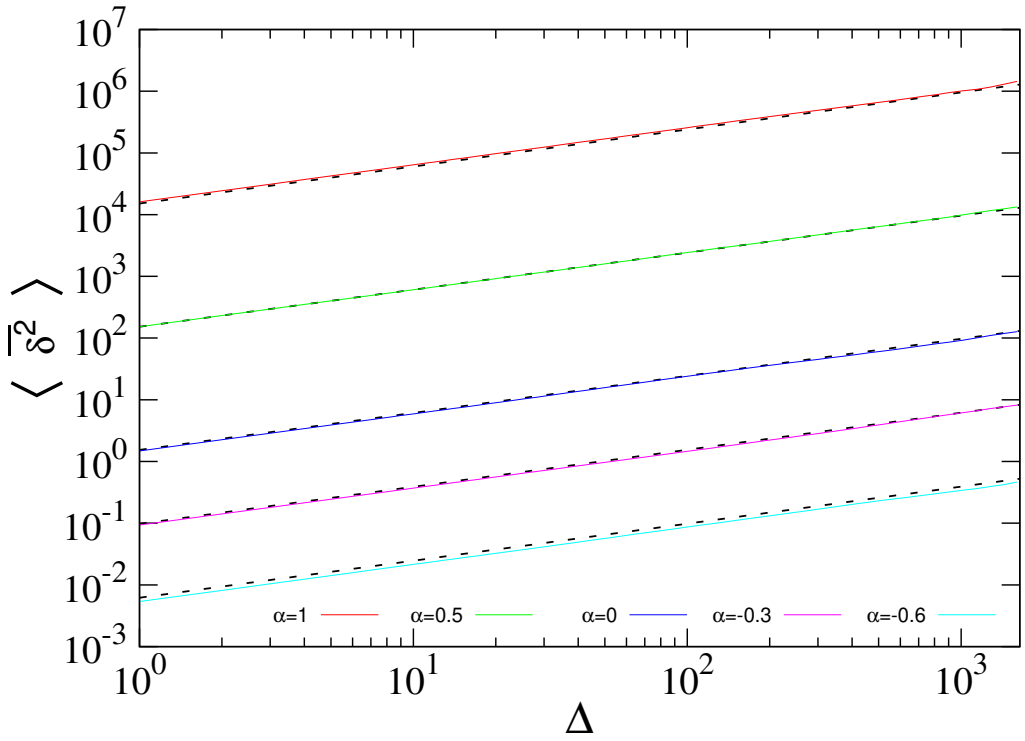


Figure 5.7: EATAMSD of the extended ggBm as a function of the lag time Δ. Dashed black lines represent the theoretical asymptotic behaviour (5.23) for $t_a \gg T \gg \Delta$. Parameters: $\beta = H = 0.3$, $t_a = 9830$, $T = 1638$, $N = 10^3$.

5.7 Derivation of asymptotic EAMSD and EATAMSD of ggBm with aging

We can calculate the EAMSD of (5.19):

$$\begin{aligned}\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a &= \left\langle [X_{\alpha,\beta,H}(t_a+t) - X_{\alpha,\beta,H}(t_a)]^2 \right\rangle \\ &= \left\langle X^2_{\alpha,\beta,H}(t_a+t)\right\rangle + \left\langle X^2_{\alpha,\beta,H}(t_a)\right\rangle - 2\left\langle X_{\alpha,\beta,H}(t_a+t)X_{\alpha,\beta,H}(t_a)\right\rangle \\ &= \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\left[(t_a+t)^{\alpha+2H} + t_a^{\alpha+2H} - (t_a+t)^{\alpha/2}t_a^{\alpha/2}\left((t_a+t)^{2H} + t_a^{2H} - t^{2H}\right)\right].\end{aligned} \tag{5.24}$$

Knowing that

$$(\tau+\Delta)^a \sim \tau^a + a\Delta\tau^{a-1}, \quad \tau \gg \Delta, \tag{5.25}$$

it is possible to obtain the behaviour when $t \gg t_a$,

$$\begin{aligned}\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a &\sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\Big[t^{\alpha+2H} + (\alpha+2H)t_a t^{\alpha+2H-1} + t_a^{\alpha+2H} \\ &\qquad - \left(t^{\alpha/2} + \frac{\alpha}{2}t_a t^{\alpha/2-1}\right) t_a^{\alpha/2}\left(t^{2H} + 2Ht_a t^{2H-1} + t_a^{2H} - t^{2H}\right)\Big] \\ &\sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\left[t^{\alpha+2H} + 2Ht_a^{\alpha/2+1}t^{\alpha/2+2H-1}\right]. \end{aligned} \tag{5.26}$$

In our case, $\alpha > -1$, $H \in (0,1)$, it results

$$\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a \sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{\alpha+2H}, \qquad t \gg t_a. \tag{5.27}$$

For $t_a \gg t$,

$$\begin{aligned}\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a &\sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\Big[t_a^{\alpha+2H} + (\alpha+2H)t\, t_a^{\alpha+2H-1} + t_a^{\alpha+2H} \\ &\qquad - \left(t_a^{\alpha/2} + \frac{\alpha}{2}t\, t_a^{\alpha/2-1}\right) t_a^{\alpha/2}\left(t_a^{2H} + 2Ht\, t_a^{2H-1} + t_a^{2H} - t^{2H}\right)\Big] \\ &= \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\left[t^{2H}t_a^{\alpha} - \alpha H t^2 t_a^{\alpha+2H-2} + \frac{\alpha}{2}t^{2H+1}t_a^{\alpha-1}\right] \end{aligned} \tag{5.28}$$

Using that $H \in (0,1)$, it completes the result (5.22)

$$\left\langle X^2_{\alpha,\beta,H}(t)\right\rangle_a \sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)} t^{2H}t_a^{\alpha}, \qquad t_a \gg t. \tag{5.29}$$

Using this last expression, it is also possible to calculate the asymptotic behaviour of the EATAMSD for $T \gg \Delta$. If $t_a \gg T \gg \Delta$,

$$\begin{aligned}\left\langle \overline{\delta^2_a(\Delta)}\right\rangle &= \frac{1}{T-\Delta}\int_{t_a}^{t_a+T-\Delta} \left\langle [X(\tau+\Delta) - X(\tau)]^2\right\rangle d\tau \\ &\sim \frac{\sigma^2 V_H \Delta^{2H}}{\Gamma(1+\beta)T}\int_{t_a}^{t_a+T} \tau^{\alpha}\, d\tau \\ &= \frac{\sigma^2 V_H \Delta^{2H}}{(\alpha+1)\Gamma(1+\beta)T}\left[(t_a+T)^{\alpha+1} - t_a^{\alpha+1}\right] \\ &\sim \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\Delta^{2H}t_a^{\alpha}. \end{aligned} \tag{5.30}$$

If $T \gg \Delta, t_a$,

$$\begin{aligned}\left\langle \overline{\delta^2_a(\Delta)}\right\rangle &= \frac{1}{T-\Delta}\int_{t_a}^{t_a+T-\Delta} \left\langle [X(\tau+\Delta) - X(\tau)]^2\right\rangle d\tau \\ &= \frac{1}{T-\Delta}\left[\int_{t_a}^{C} \left\langle [X(\tau+\Delta) - X(\tau)]^2\right\rangle d\tau + \frac{\sigma^2 V_H}{\Gamma(1+\beta)}\Delta^{2H}\int_{C}^{t_a+T-\Delta} \tau^{\alpha}\, d\tau\right], \end{aligned} \tag{5.31}$$

for a constant C that satisfies $\Delta \ll C < T + t_a - \Delta$. We can now write

$$\left\langle \overline{\delta_a^2(\Delta)} \right\rangle \sim \frac{\sigma^2 V_H \Delta^{2H}}{\Gamma(1+\beta)T} \left(A_C + \int_C^T \tau^\alpha \, d\tau \right), \tag{5.32}$$

for some constant A_C with respect to T. If $\alpha > -1$,

$$\begin{aligned} \left\langle \overline{\delta_a^2(\Delta)} \right\rangle &\sim \frac{\sigma^2 V_H \Delta^{2H}}{(\alpha+1)\Gamma(1+\beta)T} \left(B_C + T^{\alpha+1} \right) \\ &\sim \frac{\sigma^2 V_H \Delta^{2H}}{(\alpha+1)\Gamma(1+\beta)} T^\alpha. \end{aligned} \tag{5.33}$$

Chapter 6

Results Part 2: Crossover from anomalous to normal diffusion

> Madness and insanity are two terms that are so vague and relative that you can't really apportion proper values to them. The only thing I can think of that has any use it functional and dysfunctional. Are you working as well? In which case, it doesn't matter if you are mad.
>
> Alan Moore, V for Vendetta

6.1 Tempered superdiffusive fBm

Let us consider a regular overdamped Langevin equation of a particle in a viscous medium under the influence of a force [80, 94]

$$\frac{dx(t)}{dt} := v(t) = \frac{\xi(t)}{m\eta}, \quad x(0) = 0, \tag{6.1}$$

where $\xi(t)$ is a stochastic force that will be a stationary Gaussian noise with zero mean, $x(t)$ is the position of the particle, $v(t)$ its velocity, m its mass and η is the friction coefficient. Then, the velocity autocorrelation function satisfies

$$\langle v(t)v(t+\tau)\rangle = \langle v^2\rangle_\tau, \quad \tau \geq 0. \tag{6.2}$$

It holds that

$$\langle x^2(t)\rangle = 2\int_0^t d\tau (t-\tau)\langle v^2\rangle_\tau \tag{6.3}$$

so we can infer that if $\int_0^\infty d\tau\langle v^2\rangle_\tau$ is finite, then

$$\langle x^2(t)\rangle \sim 2t\int_0^\infty d\tau\langle v^2\rangle_\tau, \quad \text{as } t \to \infty, \tag{6.4}$$

being diffusion asymptotically normal. On the other hand, if $\int_0^\infty d\tau \langle v^2 \rangle_\tau$ is infinity or zero, one should expect anomalous diffusion at long times. If we choose

$$\langle v^2 \rangle_\tau = \frac{D_H}{\Gamma(2H-1)} \tau^{2H-2}, \tag{6.5}$$

where $1/2 \leq H < 1$, it is the case of superdiffusive fractional Brownian motion (note that here, different constants has been used with respect to (4.46)). In fact, the MSD is

$$\langle x^2(t) \rangle = \frac{2D_H}{\Gamma(2H+1)} t^{2H}. \tag{6.6}$$

6.1.1 Exponentially truncated fractional Gaussian noise

Let us now consider the following exponential tempering for $\langle v^2 \rangle_\tau$ (which is proportional to $\langle \xi^2 \rangle_\tau$)

$$\langle v^2 \rangle_\tau = \frac{D_H}{\Gamma(2H-1)} \tau^{2H-2} e^{-\tau/\tau_\star}, \quad \tau, \tau_\star > 0, \tag{6.7}$$

where $\tau_\star$ is a characteristic crossover time scale. Now $\int_0^\infty d\tau \langle v^2 \rangle_\tau = D_H \tau_\star^{2H-1}$ (finite) so we have normal diffusion at long times. The exact MSD is

$$\langle x^2(t) \rangle = \frac{2D_H \tau_\star^{2H}}{\Gamma(2H-1)} \left[\frac{t}{\tau_\star} \gamma\left(2H-1, t/\tau_\star\right) - \gamma\left(2H, t/\tau_\star\right) \right] \tag{6.8}$$

where $\gamma(a,z) = \int_0^z t^{a-1} e^{-t} dt$ is the incomplete γ-function. It behaves as (6.6) at short times

$$\langle x^2(t) \rangle \sim \frac{2D_H}{\Gamma(2H+1)} t^{2H}, \quad t \ll \tau_\star, \tag{6.9}$$

and at long times

$$\langle x^2(t) \rangle \sim 2D_H \tau_\star^{2H-1} t, \quad t \gg \tau_\star. \tag{6.10}$$

This crossover behaviour is plotted in Figure 6.1.

6.1.2 Power-law truncated fractional Gaussian noise

Now we use a softer power-law trunctation

$$\langle v^2 \rangle_\tau = \frac{D_H}{\Gamma(2H-1)} \tau^{2H-2} \left(1 + \frac{\tau}{\tau_\star}\right)^{-\mu}, \quad \tau, \mu > 0. \tag{6.11}$$

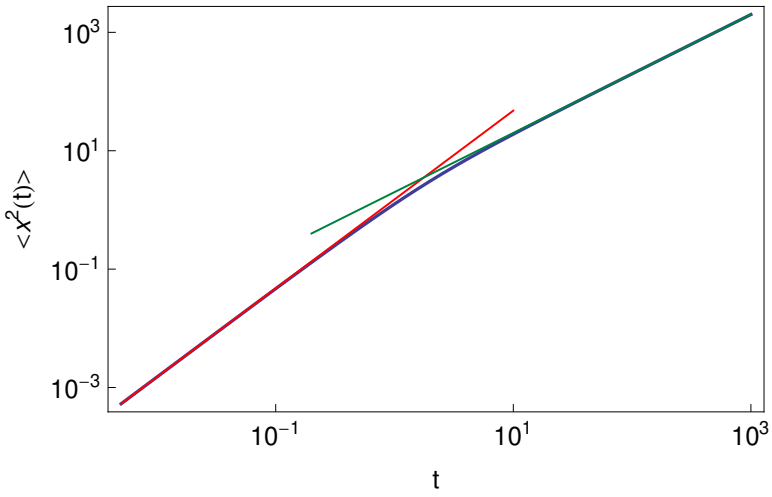


Figure 6.1: MSD (6.6) of overdamped Langevin equation with exponentially truncated fGn for $H = 3/4$, $D_H = 1$ and $\tau_\star = 1$ (blue line). Short and long time asymptotics, (6.9) (red line) and (6.10) (green line), respectively, are also plotted.

The MSD can be written as

$$\langle x^2(t)\rangle = \frac{2D_H t^{2H}}{\Gamma(2H-1)}\left[\frac{1}{2H-1}\,{}_2F_1(\mu, 2H-1; 2H; -t/\tau_\star) - \frac{1}{2H}\,{}_2F_1(\mu, 2H; 2H+1; -t/\tau_\star)\right], \tag{6.12}$$

where ${}_2F_1$ is a hypergeometric function. For $t \ll \tau_\star$ we recover the MSD of untruncated fractional Brownian motion (6.6). At long times $t \gg \tau_\star$, we can distinguish several cases:

Weak power-law truncation ($0 < \mu < 2H - 1 < 1$)

In this case, the MSD behaves as

$$\langle x^2(t)\rangle \sim \frac{2D_H \tau_\star^\mu}{(2H-\mu)(2H-1-\mu)\Gamma(2H-1)} t^{2H-\mu}, \quad t \gg \tau_\star. \tag{6.13}$$

Strong power-law truncation ($\mu > 2H - 1 > 0$)

In this situation, it is obtained

$$\langle x^2(t)\rangle \sim \frac{2D_H \Gamma(\mu - 2H + 1)\tau_\star^{2H-1} t}{\Gamma(\mu)}, \quad t \gg \tau_\star. \tag{6.14}$$

Borderline case ($0 < \mu = 2H - 1 < 1$)

The MSD of this case is

$$\langle x^2(t)\rangle \sim \frac{2D_H \tau_\star^{2H-1}}{\Gamma(2H-1)} t \ln(t/\tau_\star), \quad t \gg \tau_\star. \tag{6.15}$$

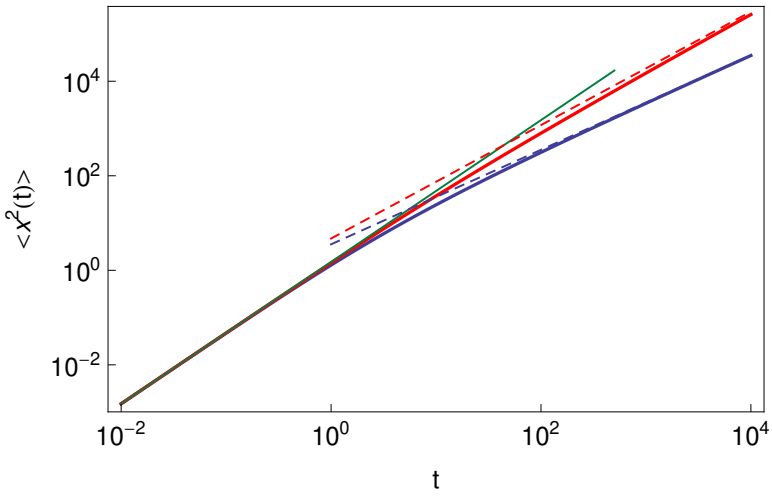


Figure 6.2: MSD (6.12) of overdamped Langevin equation with power-law truncated fGn for $H = 3/4$, $D_H = 1$ and $\tau_\star = 1$. The red solid line represents the curve for $\mu = 0.3$ (weak power-law truncation) and the blue line for $\mu = 1$ (strong power-law truncation). The dashed lines correspond to the long time asymptotics of both curves. The green line depicts the untruncated case.

In Figure 6.2 the MSD of weak and strong power-law truncation cases are shown.

6.2 Tempered subdiffusive generalised Langevin equation motion

Let us now consider the overdamped generalised Langevin equation [45, 66, 99]

$$m \int_0^t \gamma_H(t - t') \frac{dx(t')}{dt'} dt' = \xi(t), \quad x(0) = 0, \tag{6.16}$$

for a particle with mass m moving in a viscous medium that is characterized by a friction kernel $\gamma_H(t)$. $\xi(t)$ is a Gaussian noise that we will modify as in previous section. Now, we require in addition the Kubo-Zwanzig fluctuations dissipation relation [45, 99]

$$\langle \xi^2 \rangle_\tau = k_B T m \gamma_H(\tau). \tag{6.17}$$

Using Laplace transformation procedures, it is obtained that

$$\langle x^2(t) \rangle \sim \frac{2k_B T}{m \int_0^\infty \gamma_H(\tau) d\tau} t, \quad \text{as } t \to \infty. \tag{6.18}$$

Equivalently to previous section, anomalous diffusion is expected when $\int_0^\infty \gamma_H(\tau) d\tau$ is either zero or infinity. By choosing

$$\gamma_H(\tau) = \frac{\Gamma_H}{\Gamma(2H - 1)} \tau^{2H-2}, \tag{6.19}$$

with $1/2 \leq H < 1$, it is obtained that

$$\langle x^2(t)\rangle = \frac{1}{\Gamma(3-2H)}\frac{2k_BT}{m\Gamma_H}t^{2-2H}, \tag{6.20}$$

$$\langle v^2\rangle_\tau = -\frac{\sin(\pi[2H-1])\Gamma(2H)}{\pi}\frac{k_BT}{m\Gamma_H}|\tau|^{-2H}. \tag{6.21}$$

6.2.1 Exponentially truncated fractional Gaussian noise

Now we choose

$$\gamma_H(\tau) = \frac{\Gamma_H}{\Gamma(2H-1)}\tau^{2H-2}e^{-\tau/\tau_\star} \tag{6.22}$$

that arrives to

$$\langle x^2(t)\rangle = \frac{2k_BT}{m\Gamma_H}t^{2-2H}E^{1-2H}_{1,3-2H}(-t/\tau_\star) \tag{6.23}$$

where $E^\delta_{\alpha,\beta}(z)$ is the three-parameters Mittag-Leffler function. The asymptotic behaviour at short times is the same than the one for the untruncated noise (6.20). At long times,

$$\langle x^2(t)\rangle \sim \frac{2k_BT}{m\Gamma_H\tau_\star^{2H-1}}t, \quad t \gg \tau_\star. \tag{6.24}$$

The MSD and the asymptotic behaviours can be seen in Figure 6.3. The velocity autocorrelation function is, for $\tau > 0$,

$$\langle v^2\rangle_\tau = -\frac{\sin(\pi[2H-1])\Gamma(2H)}{\pi}\frac{k_BT}{m\Gamma_H}\tau^{-2H}e^{-\tau/\tau_\star}. \tag{6.25}$$

6.2.2 Power-law truncated fractional noise

Let us now use the following friction kernel

$$\gamma_H(\tau) = \frac{\Gamma_H}{\Gamma(2H-1)}\tau^{2H-2}\left(1+\frac{\tau}{\tau_\star}\right)^{-\mu}, \quad \tau,\mu > 0. \tag{6.26}$$

In this case, the MSD and the velocity autocorrelation function must be expressed in terms of the inverse Laplace transform $\mathcal{L}_s^{-1}$. In fact,

$$\langle x^2(t)\rangle = \frac{2k_BT}{m\Gamma_H\tau_\star^{2H-1}}\,g(t), \tag{6.27}$$

$$\langle v^2\rangle_\tau = \frac{k_BT}{m\Gamma_H\tau_\star^{2H-1}}\,\frac{d^2}{d\tau^2}g(\tau), \tag{6.28}$$

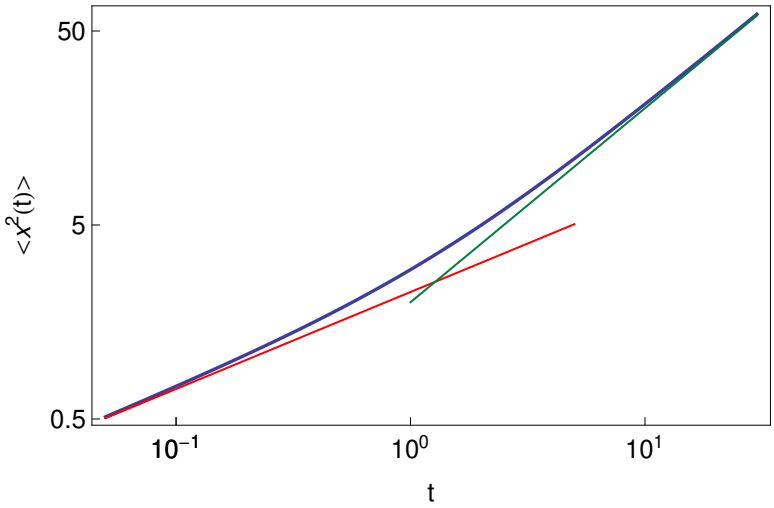


Figure 6.3: MSD (6.23) for the overdamped generalised Langevin equation with exponentially tempered fGn with parameters $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 1$ (blue line). The short and long time asymptotics are shown too (read and green lines, respectively).

where

$$g(t) = \mathcal{L}_s^{-1}\left\{\frac{1}{s^2 U(2H-1, 2H-\mu; s\tau_\star)}\right\}(t), \tag{6.29}$$

being $U(a, b; z)$ the Tricomi hypergeometric function. At short times $t, \tau \ll \tau_\star$ the MSD and the velocity autocorrelation function for the untruncated kernel are recovered. At long times, we consider two possibilities:

Weak power-law truncation ($0 < \mu < 2H - 1 < 1$)

In this case, the long time behavior is

$$\langle x^2(t)\rangle \sim \frac{\Gamma(2H-1)}{\Gamma(2H-\mu-1)\Gamma(\mu+3-2H)} \frac{2k_BT}{m\Gamma_H\tau_\star^\mu}\, t^{\mu+2-2H}, \quad t \gg \tau_\star, \tag{6.30}$$

$$\langle v^2\rangle_\tau \sim -C\frac{k_BT}{m\Gamma_H\tau_\star^\mu}\frac{1}{\tau^{2H-\mu}}, \quad \tau \gg \tau_\star, \tag{6.31}$$

where $C = (2H-\mu-1)\pi^{-1}\sin(\pi[2H-\mu-1])\Gamma(2H-1)$ is a positive constant.

Strong power-law truncation ($\mu > 2H - 1 > 0$)

In this case, MSD behaves as

$$\langle x^2(t)\rangle \sim \frac{\Gamma(\mu)}{\Gamma(\mu+1-2H)} \frac{2k_BT}{m\Gamma_H\tau_\star^{2H-1}}\, t, \quad t \gg \tau_\star. \tag{6.32}$$

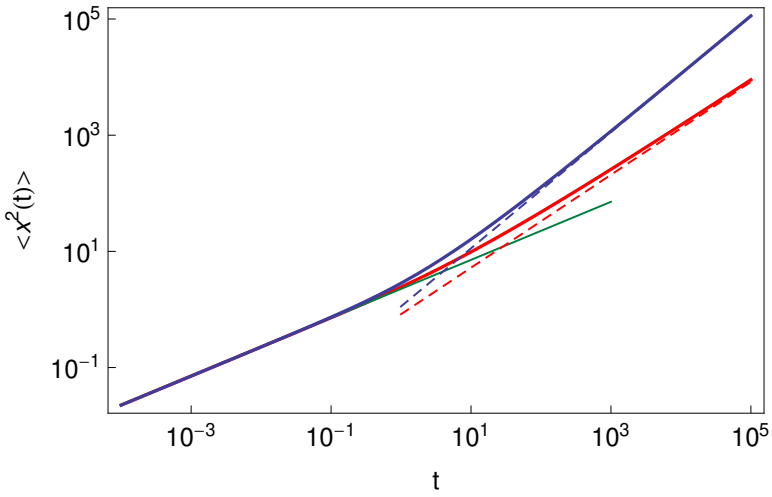


Figure 6.4: MSD (6.27) for the overdamped generalised Langevin equation with power-law truncated fGn. Common parameters are $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 1$. The red solid line shows the weak power-law truncation ($\mu = 0.3$) while the blue solid line displays the strong power-law truncation ($\mu = 1$). The asymptotics at long times are represented by dashed lines. The untruncted case is also plotted (solid green line).

For simplicity, assuming that $\mu + 1 - 2H \neq n \in \mathbb{N}$, it is obtained

$$\langle v^2 \rangle_\tau \sim -C \frac{k_BT}{m\Gamma_H \tau_\star^{2H-1}} \frac{1}{\tau^{\mu+2-2H}}, \quad \tau \gg \tau_\star, \tag{6.33}$$

where C is a positive constant.

Borderline case ($0 < \mu = 2H - 1 < 1$)

In the borderline case, at long times, MSD reads

$$\langle x^2(t) \rangle \sim \Gamma(2H-1) \frac{2k_BT}{m\Gamma_H \tau_\star^{2H-1}} \frac{t}{\ln(t/\tau_\star)}, \quad t \gg \tau_\star. \tag{6.34}$$

In Figure 6.4, the MSD for the weak and strong power-law truncation are shown.

6.3 Direct tempering of Mandelbrot's fractional Brownian motion

Meerschaert and Sabzikar considered [62] the following extension of fractional Brownian motion, the tempered fractional Brownian motion (tfBm)

$$\begin{aligned} B_{H,\lambda}(t) = &\int_{-\infty}^{0} \left[e^{-\lambda(t-t')}(t-t')^{H-1/2} - e^{-\lambda(-t')}(-t')^{H-1/2} \right] B'(t')dt' \\ &+ \int_0^t \left[e^{-\lambda(t-t')}(t-t')^{H-1/2} \right] B'(t')dt', \quad H, \lambda, t > 0. \end{aligned} \tag{6.35}$$

$B'(t)$ is white Gaussian noise with zero mean and covariance

$$\langle B'(t_1)B'(t_2)\rangle = \sigma^2\delta(t_1 - t_2). \tag{6.36}$$

In the limit $\lambda \to 0$ and $H \in (0,1)$ the fBm, as defined by Mandelbrot [58], is recovered except by a prefactor $1/\Gamma(H+1/2)$ that is not considered. This process was studied in [18] and the following results can be found there. However, they were included in the article used to recreate this Chapter for the convenience of the reader and present physical arguments for the behaviour of this model. The MSD is

$$\langle B^2_{H,\lambda}\rangle = \sigma^2 C_t^2 t^{2H} \tag{6.37}$$

where

$$C_t^2 = \left[\frac{2\Gamma(2H)}{(2\lambda t)^{2H}} - \frac{2\Gamma(H+1/2)}{\sqrt{\pi}}\frac{K_H(|\lambda t|)}{(2\lambda t)^H}\right] \tag{6.38}$$

and $K_H(z)$ is the modified Bessel function of second kind. At short times, MSD behaves as

$$\langle B^2_{H,\lambda}(t)\rangle \sim \sigma^2\Gamma^2(H+1/2)V_H t^{2H} + \frac{\sigma^2\Gamma(2H)}{2^{1+2H}(H-1)}\lambda^{2-2H}t^2, \quad t \ll \lambda^{-1}, \tag{6.39}$$

where

$$V_H = \frac{1}{\Gamma(2H+1)\sin(\pi H)}. \tag{6.40}$$

In the long time limit, the MSD converges exponentially to a constant value

$$\langle B^2_{H,\lambda}(t)\rangle \sim \sigma^2\left(\frac{2\Gamma(2H)}{(2\lambda)^{2H}} - \frac{2^{1/2-H}\Gamma(H+1/2)}{\lambda^{H+1/2}}t^{H-1/2}e^{-\lambda t}\right), \quad t \gg \lambda^{-1}. \tag{6.41}$$

In Figure 6.7 the MSD and the constant value that it approaches at long times are displayed. It is possible to define a derivative for this process using the procedure followed by Mandelbrot and van Ness

$$B'_{\lambda,H}(t;\delta) \equiv B'_{\lambda,H}(t) = \frac{1}{\delta}\left[B_{H,\lambda}(t+\delta) - B_{H,\lambda}(t)\right]. \tag{6.42}$$

which autocorrelation function results to be

$$\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\rangle = \frac{\sigma^2\Gamma(H+1/2)}{\sqrt{\pi}(2\lambda)^H\delta^2}\Big[2\tau^H K_H(|\lambda\tau|) - (\tau+\delta)^H K_H(\lambda|\tau+\delta|) - |\tau-\delta|^H K_H(\lambda|\tau-\delta|)\Big]. \tag{6.43}$$

It is antipersistent for the whole range $0 < H < 1$ and for any finite λ,

$$\int_0^\infty \left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle d\tau = 0. \tag{6.44}$$

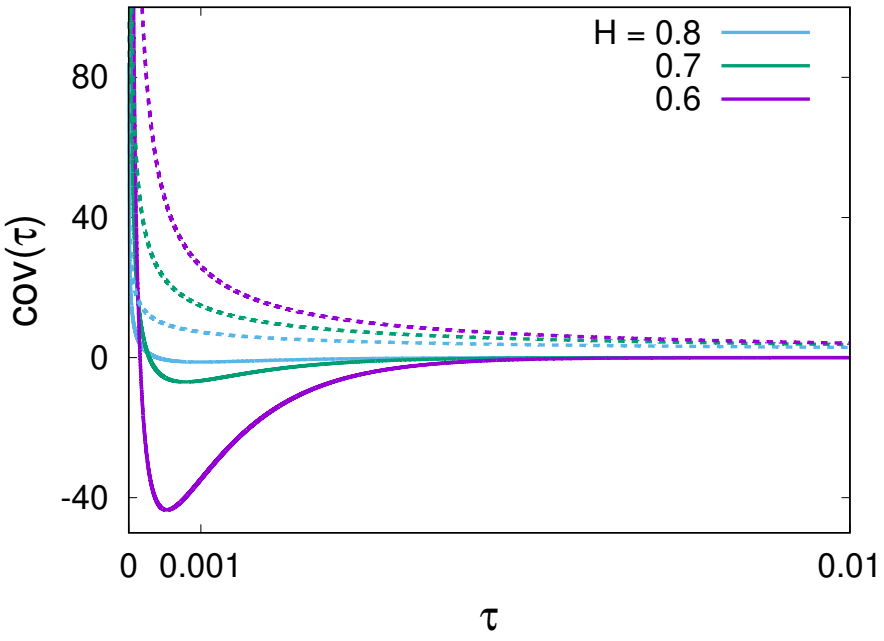


Figure 6.5: Autocorrelation function (6.43) of tempered fractional Gaussian noise (solid lines). Dashed lines show the autocorrelation function of fractional Gaussian noise. Parameters: $\lambda = 10^3$, $\delta = 10^{-5}$.

For small τ and $\delta \to 0$, it behaves as fractional Gaussian noise

$$\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\rangle \sim \sigma^2(2H-1)H\Gamma^2(H+1/2)V_H|\tau|^{2H-2} - \frac{\sigma^2\Gamma(2H)\lambda^{2-2H}}{2^{2H+1}(1-H)}, \quad \tau \ll \lambda^{-1}. \tag{6.45}$$

and, for large τ,

$$\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\rangle \sim \frac{\sigma^2\tau^{H-1/2}e^{-\lambda\tau}}{2^{H-1/2}\lambda^{H+1/2}\delta^2}\left[1-\cosh(\lambda\delta)+\sinh(\lambda\delta)\frac{(H-1/2)\delta}{\tau}\right], \quad \tau \gg \lambda^{-1}. \tag{6.46}$$

Figure 6.5 shows function (6.43) for several values of $H > 1/2$ and the corresponding curves of fractional Gaussian noise (the limit $\lambda \to 0$) for the same values of H.

6.3.1 Fractional Langevin equation with directly tempered fractional Gaussian noise

Let us consider the overdamped tempered fractional Langevin equation [18]

$$\begin{cases} \int_0^t \gamma_H(t-\tau)\frac{dx}{d\tau}d\tau = \xi(t) \\ \gamma_H(\tau) = 2\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\rangle \end{cases} \tag{6.47}$$

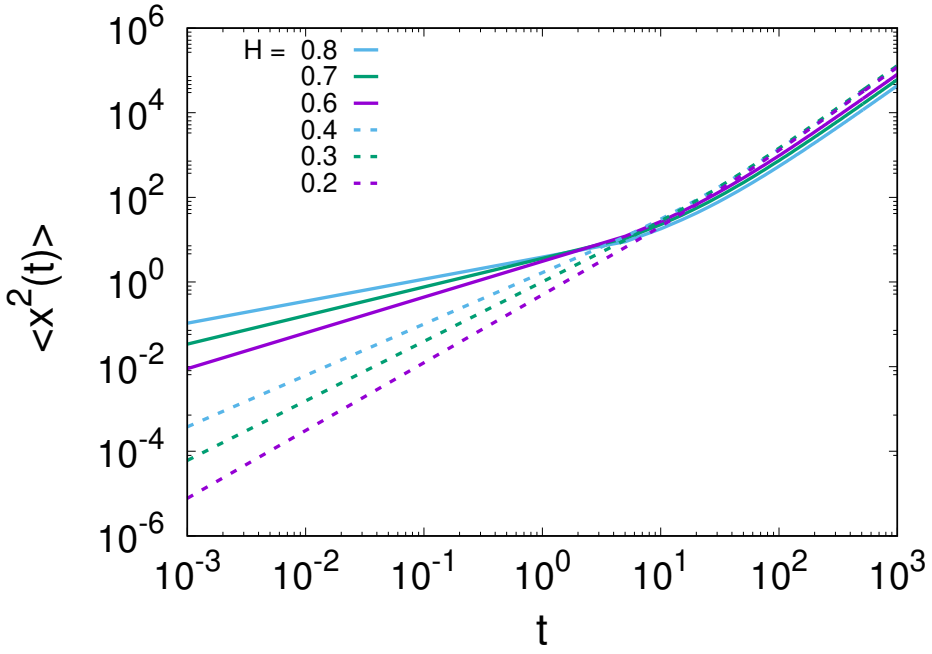


Figure 6.6: MSD of generalised Langevin with tfGn (6.47) obtained numerically. Parameter: $\lambda = 0.1$.

where the noise $\xi(t)$ is tempered fractional Gaussian noise $B'_{H,\lambda}(t)$. For the regime of short observation times, the MSD is

$$\langle x^2(t)\rangle \sim \frac{\sin(\pi H)}{\Gamma^2(H+1/2)}\frac{t^{2-2H}}{\Gamma(3-2H)}, \quad \delta \ll t \ll \lambda^{-1}. \tag{6.48}$$

For long observation times, it behaves as

$$\langle x^2(t)\rangle \sim \frac{\sqrt{\pi}\lambda^{2H}}{\Gamma(H+1/2)\Gamma(H)}t^2, \quad t \gg \lambda^{-1}. \tag{6.49}$$

In Figure 6.6 both behaviours (6.48) and (6.49) can be observed along the transition between them.

6.3.2 Ornstein-Uhlenbeck with fractional Gaussian noise

Let us consider the Ornstein-Uhlenbeck process

$$\frac{dx(t)}{dt} = -\lambda x(t) + B'_H(t), \quad x(0) = 0, \quad t > 0, \tag{6.50}$$

where $B'_H(t)$ is fractional Gaussian noise. The MSD is

$$\langle x^2(t)\rangle = \sigma^2 V_H t^{2H} e^{-\lambda t}\left[1 + \frac{\lambda t}{4H+2}(e^{\lambda t} f_H(-\lambda t) - e^{-\lambda t} f_H(\lambda t))\right], \tag{6.51}$$

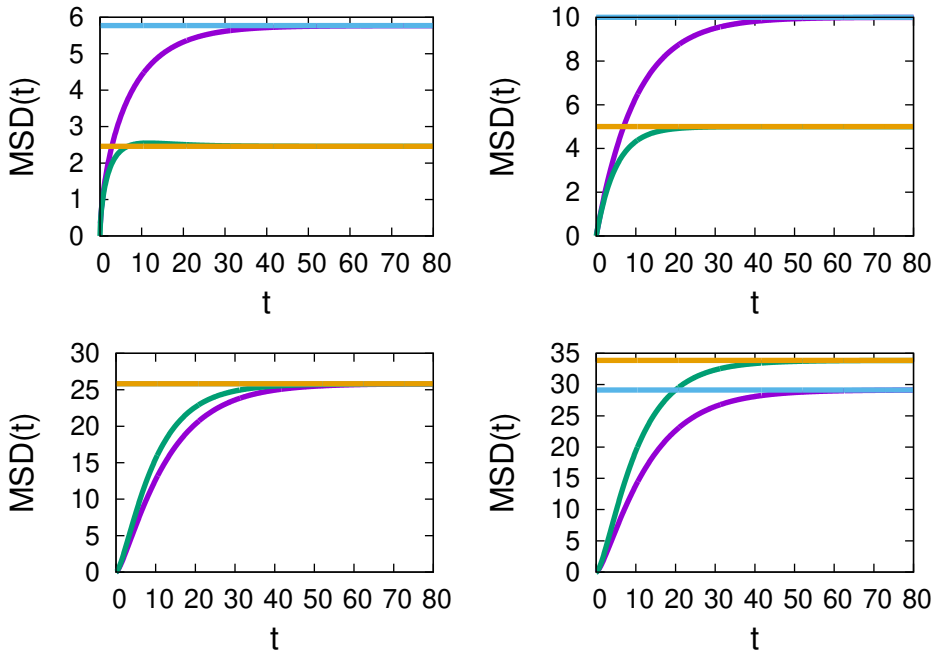


Figure 6.7: MSD of tfBm (6.37) (violet line) and asymptotic behaviour at long times (6.41) (blue line). It is also displayed the MSD of the fractional O-U process (6.51) (green line) and its long time plateau (6.54) (yellow line). The four graphs features $\sigma^2 = 1$ and $\lambda = 0.1$. The Hurst exponent is $H = 0.3$ (top left), $H = 0.5$ (top right), $H = 0.768149$ (bottom left), and $H = 0.8$ (bottom right).

where

$$f_H(x) = M(2H + 1; 2H + 2; x) \tag{6.52}$$

is Kummer's confluent hypergeometric function. For short times, MSD reads

$$\langle x^2(t)\rangle \sim \sigma^2 V_H t^{2H}(1 - \lambda t), \quad t \ll \lambda^{-1}. \tag{6.53}$$

In the long time limit, MSD converges exponentially to

$$\langle x^2(t)\rangle \sim \frac{\sigma^2}{2\sin(\pi H)\lambda^{2H}}, \quad t \gg \lambda^{-1}. \tag{6.54}$$

In Figure 6.7 the MSD and its long time limit are shown in comparison with the results known for tfBm. It can be seen the unique value of H for which both MSD reach the same plateau value.

6.4 Relation with experiments

In this section, it is shown an application to describe lipid molecule dynamics in lipid bilayer membranes. The dataset comes from all-atom Molecular Dynamics simulations regarding double layered leaves made up of short amphiphilic polymers (lipids). When they are inmersed

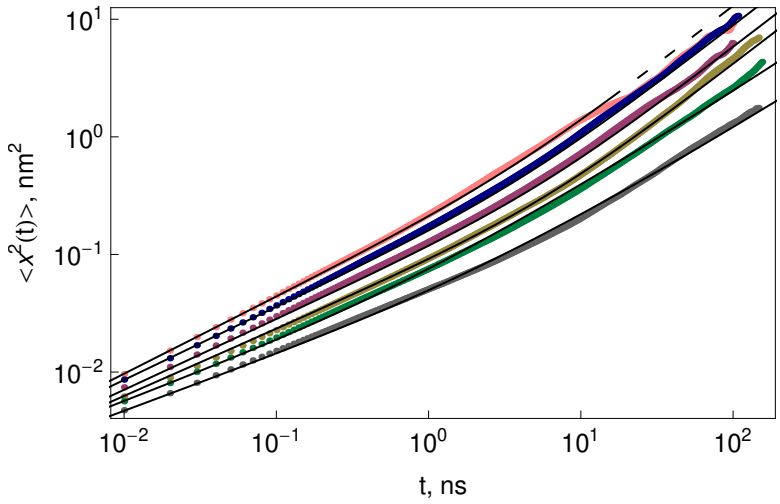


Figure 6.8: MSD of the motion of lipid molecules in a lipid bilayer model membrane at room temperature. Black solid lines are adjusted using the described generalised Langevin equation with exponentially and power-law truncated noise. The article that details this Chapter describes the parameters used for the fit. Data from Matti Javanainen, University of Helsinki.

in water, the hydrophilic head groups are in contact with the ambient (water) while the hydrophobic tail groups are not. At room temperature, the membranes form a disordered structure where membrane proteins may be embedded [95]. For such systems, simulations try to show the diffusion driven by temperature of lipids and proteins. The dynamics influence the biological properties of the membrane [95]. In Fig. 6.8, we can see the simulation results of a liquid disordered and liquid ordered state of a lipid bilayer membrane. The system is described in detail in [40]. It can be seen that diffusion changes from subdiffusion to normal diffusion and the time scale of the crossover is around 10 ns. The lipids are well described by a generalised Langevin equation featuring viscoelastic diffusion with power-law noise. It can be seen that our models with exponentially and power-law truncated noise fit very well the simulation data.

6.5 Table of main results

System	$\langle \xi^2 \rangle_\tau$	$\langle x^2(t) \rangle$, $t \to 0$	$\langle x^2(t) \rangle$, $t \to \infty$
$v(t) \propto \xi(t)$	$\propto \tau^{2H-2}$	$\propto t^{2H}$	$\propto t^{2H}$
	$\propto \tau^{2H-2} e^{-\tau/\tau_\star}$		$\propto t$
	$\propto \tau^{2H-2}(1+\tau/\tau_\star)^{-\mu}$		$\propto t^{2H-\mu}$ $(0 < \mu < 2H-1 < 1)$ $\propto t$ $(0 < 2H-1 < \mu)$ $\propto t \ln t$ $(0 < 2H-1 = \mu < 1)$
$\int_0^t \gamma_H(t-t')v(t')dt' = \xi(t)$ $\langle \xi^2 \rangle_\tau \propto \gamma_H(\tau)$	$\propto \tau^{2H-2}$	$\propto t^{2-2H}$	$\propto t^{2-2H}$
	$\propto \tau^{2H-2} e^{-\tau/\tau_\star}$		$\propto t$
	$\propto \tau^{2H-2}(1+\tau/\tau_\star)^{-\mu}$		$\propto t^{\mu+2-2H}$ $(0 < \mu < 2H-1 < 1)$ $\propto t$ $(0 < 2H-1 < \mu)$ $\propto t/\ln t$ $(0 < 2H-1 = \mu < 1)$
	$\left\langle B'_{H,\lambda}(t) B'_{H,\lambda}(t+\tau) \right\rangle$		$\propto t^2$
$x(t) = B_{H,\lambda}(t)$		$\propto t^{2H}$	const
$v(t) = -\lambda x(t) + \xi(t)$	$\langle B'^2_H \rangle_\tau \propto \tau^{2H-2}$	$\propto t^{2H}$	const

Chapter 7

Results Part 3: Stochastic spatial models in Ecology

> Desire, which has been the driving force in man, has created a great many pleasant and useful things; desire also, in man's relationships, has created a great many problems and turmoil and misery.
>
> Jiddu Krishnamurti,
> Krishnamurti to Himself

The article corresponding to this Chapter is a review. Because of that, original results come from different sources as cited.

In contrast to previous results, we are interested mainly in 2D.

7.1 Results concerning the voter model

7.1.1 β-diversity

The β-diversity for the voter model with speciation and NN dispersal is [2].

$$F(r) = c\frac{\kappa^{D-2}}{(2\pi)^{D/2}}(\kappa r)^{(2-D)/2}K_{(2-D)/2}(\kappa r), \tag{7.1}$$

where D is the number of dimensions and $\kappa^2 = 2D\nu/a^2$, being a the lattice spacing and K_z is the modified Bessel function of second kind of order z. The constant c is given by the condition $\int_{r<a} d^D r\, F(\vec{r}) = 1$.

Because of isotropy, β-diversity is only function of $r = |\vec{r}|$. It should be noticed that (7.1) is continuous in r, while the voter model is defined in a lattice. The validity of that expression is $r \gg a$ [17]. Expression (7.1) is also valid for a general dispersal kernel when distances are larger than the kernel range. In two dimensions, (7.1) displays a logarithmic decay followed by an exponential falloff. The logarithmic decay appears as a consequence of setting $D = 2$, the critical dimension of the model. On the other hand, the critical dimension allows a large biodiversity when ν is very small. It is remarkable that β-diversity measured in forests in Central and South America is compatible with the logarithmic decay for large distances [21].

7.1.2 SAR

Empirically, SAR often describes three different regimes [38, 77, 81]. At small areas A, the number of species $S(A)$ increases linearly with the sampled areas and a similar steep is shown

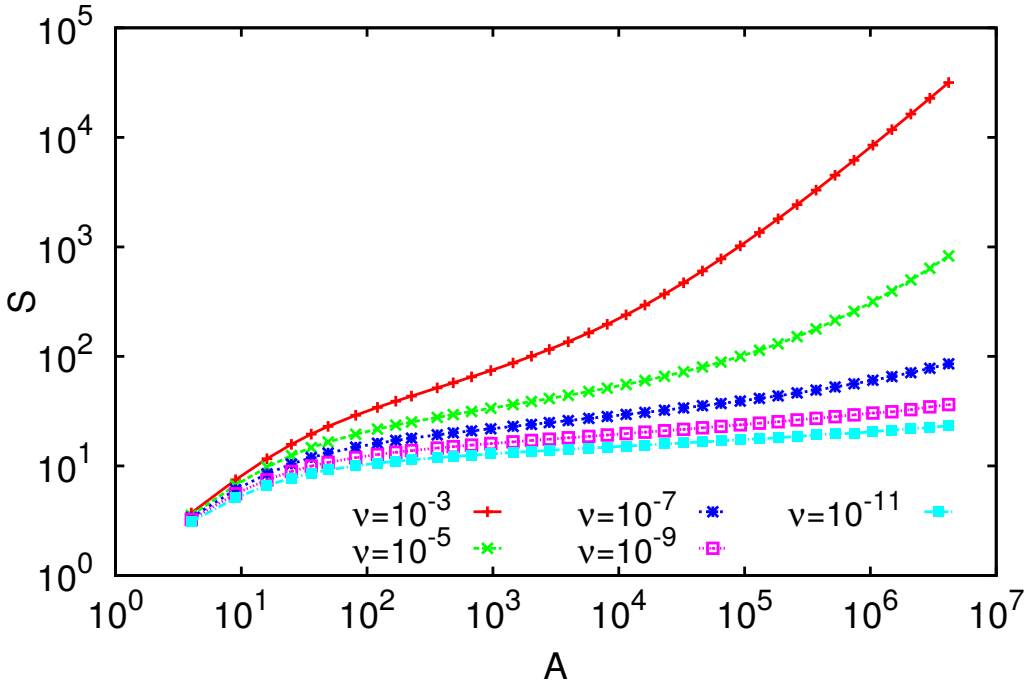


Figure 7.1: Species Area Relationship (SAR) as a function of the sampled area A for different speciation rates ν (voter model with speciation in 2 dimensions). Simulations done in [74] by using a uniform square dispersal kernel of side $K = 7$. A triphasic shape is displayed for larger ν.

at large scales. Instead, intermediate scales are described by a sublinear growth, well approximated by a power law $S(A) \sim A^z, z < 1$. It has been proposed a logarithmic fit too for that scale, $S(A) = C \ln A$.

Simulations of the voter model with speciation yields quite similar SARs to the empirical ones (see Figure 7.1). The initial regime is basically determined by the dispersal range K. It is noticed that for areas larger than K^2, a sublinear regime is obtained and for even larger areas, according to a crossover depending on ν, curves are steeper again.

Some theoretical expressions obtained for the SAR in the case of the voter model with speciation can be found in the article, along the intention to fit it to the power law form $S(A) \sim A^z$. SAR's complex dependence on A is a characteristic of the critical dimension $D = 2$.

7.1.3 SAD

For spatially explicit models such as the voter model with speciation, computing the SAD is a difficult problem. According to [2, 98], the SAD can be assumed to have the following scaling form based on standard finite-size scaling

$$P(n; A, \nu) = n^{-\beta} \Psi(n\nu^{\alpha}, A\nu^{D/2}), \tag{7.2}$$

where α and β remain unknown. It is considered the general case where $A = L^D$, being L the linear size of the sample. In models with long-range, non-diffusive dispersal [82] the scaling can be different.

In (7.2), it is described a distribution which is a power-law on n up to a scale defined by the function Ψ, depending on dimensionless products of n, ν and A. That is expected for $D = 1$ and $D \geq 3$, but not at the critical dimension $D = 2$, where logarithmic corrections should contribute to the power-law.

In order to include the logarithms, it has been developed [97] the following scaling relation

$$P(n; A) = g(A)\Psi(n/f(A)). \tag{7.3}$$

The dependence on ν was ommited since the relation was applied to empirical data (where ν is assumed to be unknown and fixed). The most important feature of this scaling relation is that f and g are general functions, which can include logarithms. They can be obtained by imposing $P(n; A)$ to be normalized and that its average value is $\langle n \rangle$. It is also used that $\Psi(x) \sim x^{-\Delta}$ with $\Delta = 1 - \epsilon$, $\epsilon \ll 1$, so it is considered a power-law on n again, but including a more general cut-off for large A. The result are the following functions (up to first order in ϵ)

$$f(A) = \langle n \rangle \ln\langle n \rangle \left[1 + \frac{\epsilon}{2} \ln\langle n \rangle\right], \tag{7.4}$$

$$g(A) = \frac{1}{\langle n \rangle \ln^2\langle n \rangle \left[1 + \frac{\epsilon}{2} \ln\langle n \rangle\right]^2}. \tag{7.5}$$

As it can be seen, they include logarithmic corrections. Zillio et al. showed that this scaling form is very good for data collected in Barro Colorado tropical forest, better than a power-law relation such as (7.2). This fact supports the idea that $\Delta \approx 1$ in tropical forests.

We tested with numerical simulations (7.3) with (7.4) and (7.5). Parameter ν was chosen in such a way that $A\nu$ is constant. Results are plotted in Figure 7.2. The small value obtained for ϵ (0.08) is consistent with the assumed small deviation from $\Delta = 1$. A similar collapse for $A\nu = 20$ reveals an even smaller value $\epsilon = 0.069$. It was verified that using $\epsilon = 0$ or removing the logarithmic corrections, the collapses are not that good. However, with the presence of statistical fluctuations, any collapse would be convincing.

This non-standard scaling form including logarithms provides an excellent collapse in empirical data and simulations of the voter model with speciation in 2D. It is reasonable to suppose that ν is kept constant between tropical forests, so the product $A\nu$ should not be constant. Our simulations (not shown) with constant ν does present deviations from perfect collapse.

Finally, a heuristic expression for the SAD of the model has been obtained following a different approach [24, 87]. Let it be $P(x, t)$ the distribution of the abundance of a given species at time t. Approximating x as a continuous quantity, a Fokker-Planck equation can be heuristically written for the evolution of $P(x, t)$

$$\partial_t P(x, t) = \nu \partial_x [x P(x, t)] + \partial_x^2 [I(x) P(x, t)], \tag{7.6}$$

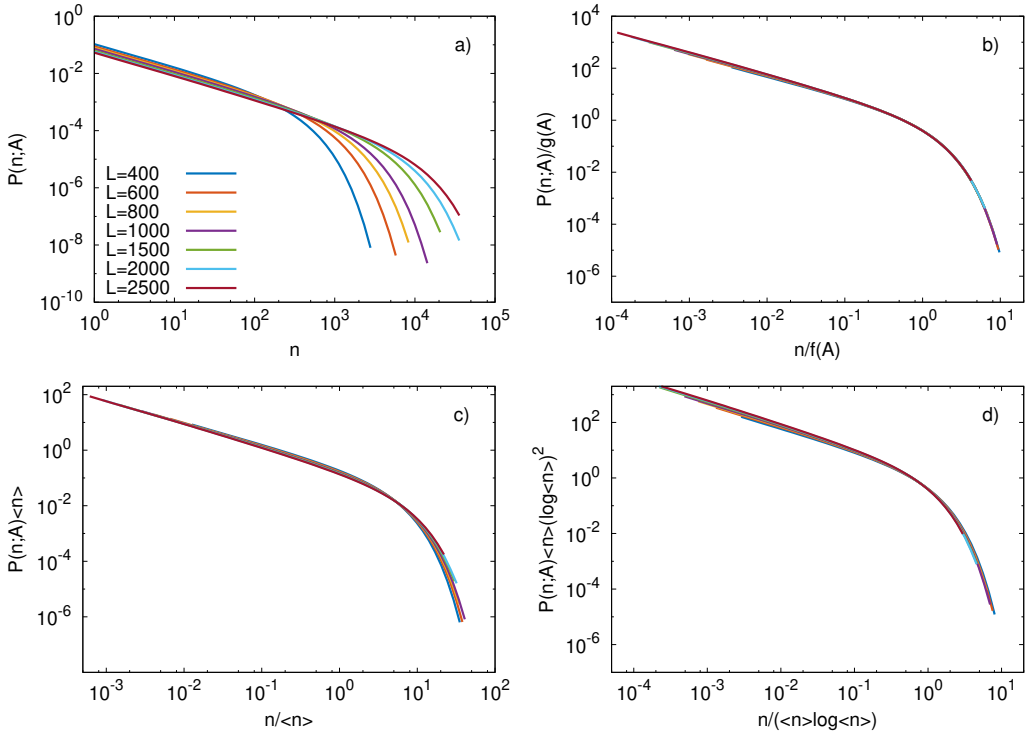


Figure 7.2: Top left: Species-Abundance Distributions (SAD) as a function of the abundance of the species n for different areas $A = L^2$ (voter model with speciation in 2 dimensions with NN dispersal kernel). The speciation rate ν is chosen in such a way that the product $A\nu$ is kept constant and equal to 200. Top right: Collapse using (7.3) and procedure described in the text with $\epsilon = 0.08$. Bottom left: Collapse without logarithmic corrections. Deviations from right collapse are noticeable. Bottom right: Collapse using (7.3) and the procedure described in the text with $\epsilon = 0$. It is also far from perfect.

where the first term in the RHS is the negative drift due to speciation, the second is the fluctuation of the population size, and $I(x)$ is the average number of interfaces of a species of size x. The crucial approximation is to not consider fluctutations of $I(x)$, which is adequate when it is a peaked function. In that situation, the dependence on the spatial dimension of the model is recap into $I(x)$. The steady-state solution is

$$P_{st}(x) = \frac{e^{-\nu \int dx \frac{x}{I(x)}}}{I(x)}. \tag{7.7}$$

In 2D, the average number of interfaces must scale as

$$I(x) = \frac{x}{1 + c \ln x} \tag{7.8}$$

where c is a non-universal constant. It is noticeable to find logarithmic terms in $I(x)$ and the fact that c plays the role of ϵ in the scaling of above. A deeper comparison of the results of the two formulations would be an interesting issue.

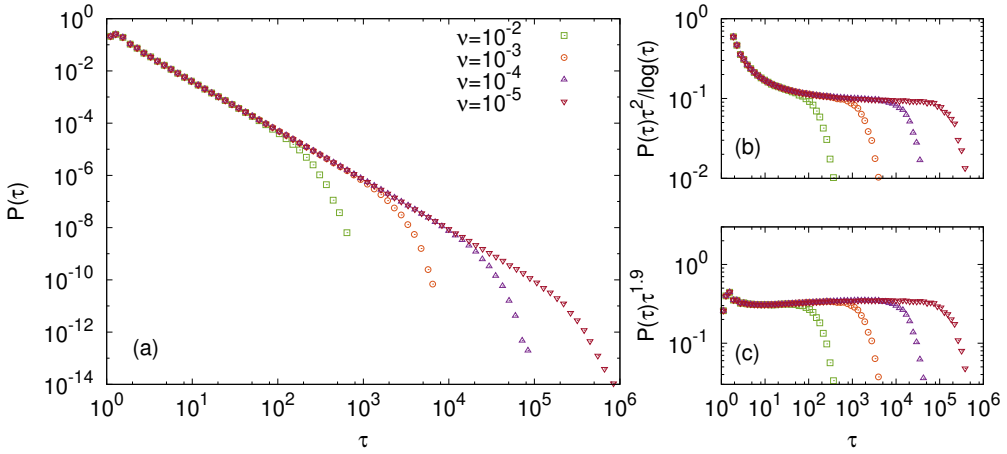


Figure 7.3: Species persistence times for voter model with speciation in 2D with NN dispersal. Left: pdf of species persistence times for different values of the speciation rate ν. Top right: pdf rescaled with the logarithmic correction. Bottom right: pdf reescaled with an effective power-law.

7.1.4 Species persistence times

Empirical measures of North-American birds and herbaceous plants show that the probability of observing persistence times τ are power-laws ($P(\tau) \sim \tau^{-1.83}$ and $P(\tau) \sim \tau^{-1.78}$, respectively), with exponential cut-offs that are area-dependent [6, 90].

The species persistence-times can be calculated for the model in the limit of vanishing ν and it is (in 2D)

$$P(\tau) \sim \frac{\ln \tau}{\tau^2}, \tag{7.9}$$

while for 1D it is

$$P(\tau) \sim \tau^{-1/2}. \tag{7.10}$$

When ν is not negligible, there are cut-offs which have the form $\exp(-\nu\tau)$ in either dimension.

Species persistence times are shown in Figure 7.3, obtained with simulations in 2D of the model. It is additionally shown a compensated plot to check the fitness to the predicted law. It is also shown that using an effective power-law, a good approximation in a broad range of scales is obtained. That is consistent with empirical results [6, 90].

7.2 Near-neutral models

In real ecosystems, neutral models are a raw approximation. In this section, it is considered a minimal model where a single parameter determines if the model remains neutral or becomes non-neutral. This model is the habitat-preference model [75] and it is a variant of the voter model.

7.2.1 Description of the system

Only two species are considered (A and B) with N_A and N_B individuals, respectively. The total number of individuals is $N = N_A + N_B$ and they are placed in a saturated square lattice, where periodic boundary conditions apply. There is an additional infinite reservoir where species A and B can migrate to the lattice. Now we impose the condition of non-neutrality allowing each site of the lattice to have a fixed preference for individuals of the species A or B, so we associate to each node a type a or b according to its preference for the individual of species A and B, respectively, when the colonization from the reservoir takes place. Mortality and dispersal are independent of the type of the site so the habitat-preference is related with the seeds and not with the individuals themselves.

7.2.2 Dynamics

Each time step, the individual of a site of the lattice randomly chosen dies. Then, a new individual is associated to that node, which comes from the reservoir with probability μ or by reproduction of one of the four nearest neighbours with probability $1 - \mu$. In both cases, the probability to colonize the empty site by those individuals which are prefered by the type of the site is multiplied by a factor γ. The exact formulas are included in the paper, also considering a global dispersal or mean-field version. Note that for $\gamma = \mu = 0$, the standard voter model is recovered (and, consequently, its neutrality).

7.2.3 Extinction times

When $\mu = 0$ (abscence of inmigration) and there is a finite number of individuals N, persistent coexistence of A and B is impossible. Instead, the system reaches eventually one of the two absorving states (monodominance of one of the species). An observable of interest is the average extintion time $\langle T_{\text{ext}} \rangle$.

The neutral case (voter model, $\gamma = 0$) has the following extintion time for large N and NN dispersal kernel [44]

$$\langle T_{\text{ext}} \rangle \sim N \ln N, \tag{7.11}$$

and the following for global dispersal [30]

$$\langle T_{\text{ext}} \rangle \sim N. \tag{7.12}$$

In general, $\gamma > 0$ leads to an increase of $\langle T_{\text{ext}} \rangle$ that takes the form

$$\langle T_{\text{ext}} \rangle \sim \exp(C(\gamma)\, N), \tag{7.13}$$

where $C(\gamma)$ is well-fitted by a power-law with exponent 1.63 (see Figure 7.4). The global dispersal version has a qualitative similar behaviour, except in the limit $\gamma = 0$ (as indicated in (7.12)) and the shape of $C(\gamma)$ [75]. The exponential dependence indicates that the model presents a stabilizing feature. For large enough N, both species coexist on realistic time scales.

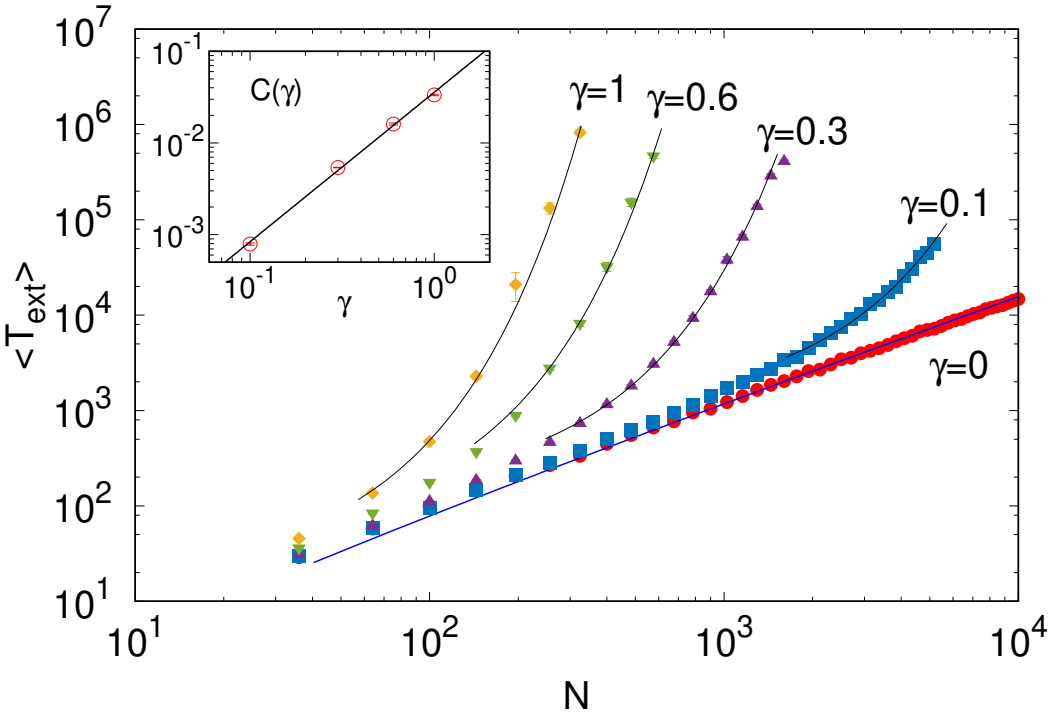


Figure 7.4: Extintion times averaged over 10^3 realizations for the habitat-preference model with nearest neighbours dispersal and $\mu = 0$. The blue curve is $N \ln N$. Black curves corresponds to the exponential fit (7.13). The points of the inset display $C(\gamma)$ and its black line is the best fit to $C(\gamma) = A\gamma^{\beta}$ ($\beta = 1.63$).

7.2.4 Coexistence

A extinct species can appear again in the case $\mu > 0$ due to immigration from the reservoir. When $1/\mu \gg \langle T_{\text{ext}} \rangle$, the recovery is unlikely, so the distribution of the population size is peaked at 0 and N for any of the species (monodominance). When $1/\mu \ll \langle T_{\text{ext}} \rangle$, extintions are rare and the distribution is peaked at $N/2$ (pure coexistence). For intermediate values, three maximum values are found for the population size distribution (0, $N/2$ and N) (mixed coexistence). These three cases are displayed in the top panels of Figure 7.5.

In the bottom panels of Figure 7.5 it can be seen the three regimes of coexistence in the N-μ parameter space. In the global dispersal version, the same qualitative panels are found, but there is no mixed coexistence regime between monodominance and pure coexistence in the case $\gamma = 0$ [75]. It can be noticed that increasing γ (habitat preference), the region of mixed coexistence expands.

7.2.5 Generalizations of the habitat-preference model

To obtain physical understanding of the different regimes in Figure 7.5, a variant of the habitat-preference model was considered in [8]. Another study analyzed the case in which only some nodes have preference for a specific species [9]. In particular, it was studied the case in which only the left column of sites of a square lattice has preference for one of the species while the right column of sites has preference for the other species, remaining the rest of

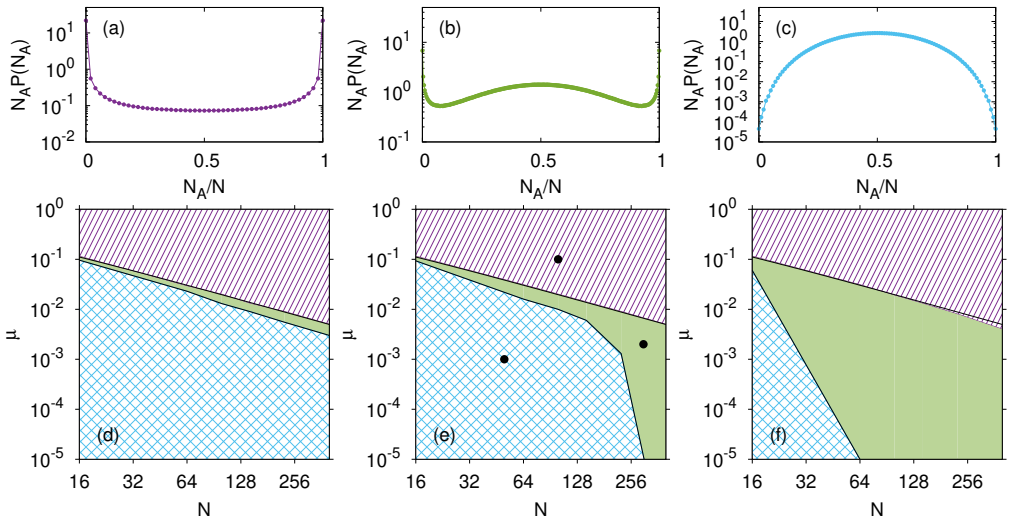


Figure 7.5: Regimes of coexistence for the habitat-preference model with NN dispersal. Top panels show the stationary distribution of the population of one of the species $P(N_A)$ for $\gamma = 0.3$ and, from left to right: $N = 50, \mu = 10^{-3}$; $N = 300, \mu = 2 \cdot 10^{-3}$; $N = 100, \mu = 10^{-3}$. Bottom panels show how the N-μ parameter space is partitioned in three regimes for different values of γ (from left to right: γ=0, 0.3 and 1). The three points are associated with the graphs in the top panel.

the nodes neutral. Results show that even mild preferences in a small fraction of nodes has as a consequence a durable coexistence. It applys also to regions very far from the biased nodes. These results have several applications since it shows that "sanctuaries" for different species can result in an improvement of coexistence of all the species, protecting biodiversity [9].

7.2.6 Temporally-dependent habitat preferences

There is a long tradition in Ecology to understand what happens when preference of the sites changes over time. Studies have looked at the consequences of changes on population growth and stability of the ecosystem [19,79]. Enviromental fluctuations can increase extinction rates, but also can foster stability by reducing interspecific competition.

Several works have studied models in which habitat-preferences are dependent on time. They predict ecological observables better than other available neutral models [25,35,87,88]. It is claimed [1] that this approach estimates more realisticaly dynamical quantities, e.g., average species persistence times, when compared with neutral models.

7.2.7 Models with density dependence

Variants of the voter model with negative density-dependence have been considered [68]. It means that individuals belonging to more abundant species have lower fitness. These models of density-dependence are symmetric, although not neutral, since all species are treated equally. Phenomena like spontaneous breakdown of that symmetry, leading to species coexistence, has been considered at mean field [7].

Chapter 8

Discussion

> There is nothing wrong with wanting pay for work, or seeking to maximize one's income, as long as one does not use means that are destructive.
>
> Richard Matthew Stallman,
> GNU Manifesto

8.1 First part

There exists a lot of anomalous diffusion models and the choice of a specific model depends on the physical situation. In addition, to determine the type of motion is usually necessary to check also additional measures, not only the MSD. In the first part we started by obtaining the observables of interest of the process known as ggBm to be able to discern whether this model can explain some experiments or not. Comparing the characteristics of ggBm with the two most popular models that reproduce anomalous diffusion, fBm and subdiffusive CTRW, we see that it has properties similar to each one. A very important remark is that, whilst the observables related with single trajectories (p-variation and TAMSD) are very similar to fBm, measures related with ensemble-averages (pdf and Ergodicity Breaking parameter) are close to CTRW. There are differences however with respect to CTRW. For example, the parameter regulating the Ergodicity Breaking parameter (and the non-ergodicity) and the order of the M-Wright/Mainardi function describing its pdf is not related with the anomalous diffusion exponent in the case of the ggBm, while it is in the case of CTRW. Also, for CTRW, non-ergodicity is a consequence of the ausence of a waiting-time scale which directly affects the TAMSD. On the contrary, ggBm is similar to fBm in the sense that it does not experience trapping and each trajectory has a self-averaging TAMSD. However, the fact that each particle experiments the media in a different way sets a non-vanishing variance for this observable and, as a consequence, non-ergodicity. This property was demonstrated for any process composed of identical ergodic processes in an heterogeneous media and the Ergodicity Breaking parameter was calculated. It would be very interesting to know how these results are related with environments displaying other kind of heterogeneity and see their relation with ergodicity.

Characteristics of ggBm has been proved to be very useful because ggBm can explain scenarios that cannot be reproduced by some of the more important models like fBm or subdiffusive CTRW. It is the case of the motion of mRNA in E. coli cells studied by Goldin and Cox where ggBm can explain all the measured observables: anomalous diffusion, fBm-like p-variation, and non-ergodicity. This way, we showed that a fractional Brownian motion in a heterogeneous media may describe this process. A very similar result was obtained recently

in [50] where data of Goldin and Cox was statistically compared with a process almost similar to ggBm, a superstatistical fBm, $Y(t) = XB_H(t)$, which random variable X is distributed as a Weibull distribution instead of the M-Wright/Mainardi distribution used in the ggBm. Despite our comparison for p-variation is exclusively qualitative, authors of that reference developed two statistical methods (one of them based on p-variation and the other on the second moment of the increments of the process) to confirm whether the motion of mRNA could be explained by the model and an afirmative answer was obtained.

Another recent paper that deals with a different ggBm is [89]. In that reference, the Gaussian process of formula (3.1) is chosen to be Brownian motion instead of fBm, and the random variable is understood as a diffusivity instead of a length scale. It can be written as $X(t) = \sqrt{2D}B(t)$, where random diffusivity D is distributed according to the generalised Gamma distribution. A non-equilibrium generalization is also proposed in which they introduce a time-dependency in the diffusivity $D = D(t)$. It is suggested to be understood as fluctuations of their disjointed regions, or also changes in the particle size, e.g., due to agglomeration and separation dynamics.

8.2 Second part

In the second part, firstly, we analysed the overdamped Langevin equation and overdamped generalised Langevin equation with different truncations for the autocorrelation function of the fractional Gaussian noise. While at short times, the MSD was the same than the case without truncation, we showed that at long times the situation changes. When exponentially truncated or strongly truncated by a power-law, normal diffusion appears at long times. These results are very important because there are experiments in which after a given correlation time, anomalous diffusion crosses over to normal diffusion. When possible, it is better to have a model that reproduces both the initial and the terminal regimes instead of two disconnected power-laws. The mentioned results are explicit and they allow to extract the crossover time. We applied one of these models to simulation results from lipid bilayer membranes where the correspondence was very good and data perfectly fitted. Many other experiments can benefit from these results since a quantitative description of the crossover was found.

In the case of the overdamped Langevin equation which correlation function of the noise has weak power-law truncation, a crossover from faster superdiffusion to slower superdiffusion was found. On the other hand, in the case of the overdamped generalised Langevin equation with the same noise that in the previous case, a crossover from slower to faster subdiffusion was obtained. Again, experiments can benefit from these models since explicit expression for the results was derived.

Secondly, we went into the process known as tempered fractional Brownian motion where an exponential tempering is done directly in the definition of the fBm. As commented in Chapter 6, these results were previously obtained by other authors but we considered important to present them for comparison and discussion. At short times, the behaviour of the MSD is like fBm except by a prefactor not considered in the definition. Surprisingly at first, the long time behaviour of the MSD is not linear as it was in the previous cases for exponential truncations, but it approaches to a constant value. As a consequence, we can think than this process is physically more suitable for a velocity than the position of a particle. We compared it with the fractional Ornstein-Uhlenbeck and both saturates at long times, so we can say that this

tempered fBm describes also a confined motion. The plateau values are dependent on H and we found the single value for which both MSD arrive to the same limit.

Another surprising result is the MSD of the fractional Langevin equation with tempered fractional Gaussian noise (the defined derivative of the tempered fBm). This model shows a crossover from subdiffusion (in agreement with previous studied equations) to a ballistic regime. We arrived to the conclussion that this last behaviour is a consequence of two facts: (i) the integral of the autocorrelation function of the noise over the entire domain is zero, and (ii) at long times this autocorrelation function exhibits an exponential decay. To show this, in the original reference of this part of the thesis, a toy model satisfying both conditions is presented and, again, it is reproduced the ballistic long time behaviour. We consider that this fact increases the understanding of the process and will help choose when this process can be useful.

8.3 Third part

With respect to community ecology, we discussed the state of the art in spatial neutral Ecology. We also considered near-neutral models and we commented them. The list of discussed points was not complete, and there are a lot of interesting issues that still need to be studied deeply. One of them is the considered speciation mechanisms. In all the models discussed in the article, the speciation process requires a single individual. However, while it is practical for modeling, it leads to speciation rates that are not equivalent to independent estimates [78]. This assumption increases the number of rare species. A process called protracted speciation has been considered for neutral models to address this problem [83]. In protracted speciation, a single generation is not enough for speciation. Instead, it becomes a gradual event that lasts several generations. In addition, there are more speciation modes [22]. It is the case of parapatric speciation, where two distant populations of the same species can give rise to different species. Different speciation mechanisms can have implications in maintaining biodiversity and ecological patterns. Computational results of models featuring these mechanisms can be useful to clarify it.

Ecological neutrality has elicited controversies because classical ecological features like niches might be not that important in structuring communities. On the other hand, neutral and non-neutral models based on niches have similar results, so it is difficult to determine [16, 60, 92]. Niche and neutral models are not exclusive, but it is not easy to distinguish the relative importance of each of them, in particular because of the big number of parameters of non-neutral models. Some progress has been done in [33], where the main aspects of neutral theory are added to the classical competition model Lotka-Volterra. To study a similar approach in a spatial model could be an interesting direction of research.

Some observables in ecology are commonly used in Statistical Physics, like β-diversity/two-point correlation function. However, other observables are not so common like SAR and SAD. Considering well-known quantities like multi-point correlation functions and measuring it in real ecosystems could be an interesting direction too.

Finally, it can be remarked that ecosystems are usually 2-dimensional, which due to the underlying diffusive behavior, is the critical dimension. It has been shown that it leads to logarithmic corrections of the scaling laws. This fact usually makes difficult the mathematical and computational analysis, even if some advances have been made in last years.

Bibliography

[1] Sandro Azaele, Simone Pigolotti, Jayanth R Banavar, and Amos Maritan. Dynamical evolution of ecosystems. *Nature*, 444(7121):926, 2006.

[2] Sandro Azaele, Samir Suweis, Jacopo Grilli, Igor Volkov, Jayanth R Banavar, and Amos Maritan. Statistical mechanics of ecological systems: Neutral theory and beyond. *Reviews of Modern Physics*, 88(3):035003, 2016.

[3] Eli Barkai. Aging in subdiffusion generated by a deterministic dynamical system. *Physical review letters*, 90(10):104101, 2003.

[4] Christian Beck. Superstatistical brownian motion. *Progress of Theoretical Physics Supplement*, 162:29–36, 2006.

[5] Christian Beck and EGD Cohen. Superstatistics. *Physica A: Statistical mechanics and its applications*, 322:267–275, 2003.

[6] Enrico Bertuzzo, Samir Suweis, Lorenzo Mari, Amos Maritan, Ignacio Rodríguez-Iturbe, and Andrea Rinaldo. Spatial effects on species persistence and implications for biodiversity. *Proceedings of the National Academy of Sciences*, 108(11):4346–4351, 2011.

[7] C. Borile, M. A. Muñoz, S. Azaele, Jayanth R. Banavar, and A. Maritan. Spontaneously broken neutral symmetry in an ecological system. *Phys. Rev. Lett.*, 109:038102, Jul 2012.

[8] Claudio Borile, Amos Maritan, and Miguel A Muñoz. The effect of quenched disorder in neutral theories. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(04):P04032, apr 2013.

[9] Claudio Borile, Daniel Molina-Garcia, Amos Maritan, and Miguel A Muñoz. Coexistence in neutral theories: interplay of criticality and mild local preferences. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(1):P01030, 2015.

[10] I Bronstein, Y Israel, E Kepten, S Mai, Yaron Shav-Tal, E Barkai, and Y Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical review letters*, 103(1):018102, 2009.

[11] S. Burov, R. Metzler, and E. Barkai. Aging and nonergodicity beyond the khinchin theorem. *Proceedings of the National Academy of Sciences*, 107(30):13228–13233, 2010.

[12] Stas Burov, Jae-Hyung Jeon, Ralf Metzler, and Eli Barkai. Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking. *Physical Chemistry Chemical Physics*, 13(5):1800–1812, 2011.

[13] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009.

[14] Massimo Cencini, Simone Pigolotti, and Miguel A Muñoz. What ecological factors shape species-area curves in neutral models? *PloS one*, 7(6):e38232, 2012.

[15] J. M. Chambers, C. L. Mallows, and B. W. Stuck. A method for simulating stable random variables. *Journal of the American Statistical Association*, 71(354):340–344, 1976.

[16] Jérôme Chave, Helene C. Muller-Landau, and Simon A. Levin. Comparing classical community models: Theoretical consequences for patterns of diversity. *The American Naturalist*, 159(1):1–23, 2002. PMID: 18707398.

[17] Jérôme Chave and Egbert G. Leigh. A spatially explicit neutral model of β-diversity in tropical forests. *Theoretical Population Biology*, 62(2):153 – 168, 2002.

[18] Yao Chen, Xudong Wang, and Weihua Deng. Localization and ballistic diffusion for the tempered fractional brownian-langevin motion. *arXiv preprint arXiv:1704.03312*, 2017.

[19] Peter L Chesson and Robert R Warner. Environmental variability promotes coexistence in lottery competitive systems. *The American Naturalist*, 117(6):923–943, 1981.

[20] Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

[21] Richard Condit, Nigel Pitman, Egbert G. Leigh, Jérôme Chave, John Terborgh, Robin B. Foster, Percy Núñez, Salomón Aguilar, Renato Valencia, Gorky Villa, Helene C. Muller-Landau, Elizabeth Losos, and Stephen P. Hubbell. Beta-diversity in tropical forest trees. *Science*, 295(5555):666–669, 2002.

[22] Jerry A Coyne and H Allen Orr. *Speciation*. Sinauer Associates, Inc, 2004.

[23] William Croft. The darwinization of linguistics. *Selection*, 3(1):75–91, 2002.

[24] Matan Danino, Yahav Shem-Tov, and Nadav M Shnerb. Spatial neutral dynamics. *arXiv preprint arXiv:1606.02837*, 2016.

[25] Matan Danino, Nadav M. Shnerb, Sandro Azaele, William E. Kunin, and David A. Kessler. The effect of environmental stochasticity on species richness in neutral communities. *Journal of Theoretical Biology*, 409:155 – 164, 2016.

[26] Weihua Deng and Eli Barkai. Ergodic properties of fractional Brownian-Langevin motion. *Physical Review E*, 79(1):011112, 2009.

[27] Ton Dieker. Simulation of fractional brownian motion. *MSc theses, University of Twente, Amsterdam, The Netherlands*, 2004.

[28] Richard Durrett and Simon A Levin. Stochastic spatial models: a user's guide to ecological applications. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 343(1305):329–350, 1994.

[29] Rick Durrett and Simon Levin. Spatial models for species-area curves. *Journal of Theoretical Biology*, 179(2):119–127, 1996.

[30] John H Gillespie. *Population genetics: a concise guide*. JHU Press, 2004.

[31] Ido Golding and Edward C Cox. Physical nature of bacterial cytoplasm. *Physical Review Letters*, 96(9):098102, 2006.

[32] David Griffeath. The basic contact processes. *Stochastic Processes and their Applications*, 11(2):151–185, 1981.

[33] Bart Haegeman and Michel Loreau. A mathematical synthesis of niche and neutral theories in community ecology. *Journal of Theoretical Biology*, 269(1):150 – 165, 2011.

[34] Y He, S Burov, R Metzler, and E Barkai. Random time-scale invariant diffusion and transport coefficients. *Physical Review Letters*, 101(5):058101, 2008.

[35] Jorge Hidalgo, Samir Suweis, and Amos Maritan. Species coexistence in a neutral dynamics with environmental noise. *Journal of Theoretical Biology*, 413:1 – 10, 2017.

[36] Haye Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in physics*, 49(7):815–958, 2000.

[37] Richard A Holley, Thomas M Liggett, et al. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability*, 3(4):643–663, 1975.

[38] Stephen P Hubbell. *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Princeton University Press, 2001.

[39] Jae-Hyung Jeon and Ralf Metzler. Fractional brownian motion and motion governed by the fractional langevin equation in confined geometries. *Phys. Rev. E*, 81:021103, Feb 2010.

[40] Jae-Hyung Jeon, Hector Martinez-Seara Monne, Matti Javanainen, and Ralf Metzler. Anomalous diffusion of phospholipids and cholesterols in a lipid bilayer and its origins. *Phys. Rev. Lett.*, 109:188103, Oct 2012.

[41] Maurice Kendall and Alan Stuart. The advanced theory of statistics. vol. 1: Distribution theory. *London: Griffin, 1977, 4th ed.*, 1977.

[42] E Kepten, I Bronshtein, and Y Garini. Ergodicity convergence test suggests telomere motion obeys fractional dynamics. *Physical Review E*, 83(4):041919, 2011.

[43] M. KIMURA. 'stepping stone' model of population. *Annual Report of the National Institute of Genetics Japan*, 3:62–63, 1953.

[44] P. L. Krapivsky. Kinetics of monomer-monomer surface catalytic reactions. *Phys. Rev. A*, 45:1067–1072, Jan 1992.

[45] R Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255, 1966.

[46] Paul Langevin. Sur la théorie du mouvement brownien. *Compt. Rendus*, 146:530–533, 1908.

[47] Simon A Levin. The problem of pattern and scale in ecology: the robert h. macarthur award lecture. *Ecology*, 73(6):1943–1967, 1992.

[48] T.M. Liggett. *Interacting particle systems*. Springer Verlag, 1985.

[49] Ariel Lubelski, Igor M Sokolov, and Joseph Klafter. Nonergodicity mimics inhomogeneity in single particle tracking. *Physical review letters*, 100(25):250602, 2008.

[50] Arleta Maćkała and Marcin Magdziarz. Statistical analysis of superstatistical fractional brownian motion and applications. *Physical Review E*, 99(1):012143, 2019.

[51] Marcin Magdziarz and Joseph Klafter. Detecting origins of subdiffusion: p-variation test for confined systems. *Physical Review E*, 82(1):011129, 2010.

[52] Marcin Magdziarz, Jakub Karol Ślęzak, and Justyna Wójcik. Estimation and testing of the hurst parameter using p-variation. *Journal of Physics A: Mathematical and Theoretical*, 46(32):325003, 2013.

[53] Marcin Magdziarz, Aleksander Weron, Krzysztof Burnecki, and Joseph Klafter. Fractional brownian motion versus the continuous-time random walk: A simple test for subdiffusive dynamics. *Physical Review Letters*, 103(18):180602, 2009.

[54] Marcin Magdziarz, Aleksander Weron, and Karina Weron. Fractional fokker-planck dynamics: Stochastic representation and computer simulation. *Phys. Rev. E*, 75:016708, Jan 2007.

[55] Francesco Mainardi, Antonio Mura, and Gianni Pagnini. The m-wright function in time-fractional diffusion processes: A tutorial survey. *International Journal of Differential Equations*, 2010, 2010.

[56] Francesco Mainardi and Gianni Pagnini. The wright functions as solutions of the time-fractional diffusion equation. *Applied Mathematics and Computation*, 141(1):51 – 62, 2003. Advanced Special Functions and Related Topics in Differential Equations, Third Melfi Workshop, Proceedings of the Melfi School on Advanced Topics in Mathematics and Physics.

[57] Francesco Mainardi, Gianni Pagnini, and RK Saxena. Fox h functions in fractional diffusion. *Journal of Computational and Applied Mathematics*, 178(1):321–331, 2005.

[58] Benoit B. Mandelbrot and John W. Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[59] Joaquin Marro and Ronald Dickman. Nonequilibrium phase transitions in lattice models. *Nonequilibrium Phase Transitions in Lattice Models, by Joaquin Marro and Ronald Dickman, pp. 344. ISBN 0521480620. Cambridge, UK: Cambridge University Press, May 1999.*, page 344, 1999.

[60] B. McGill. Strong and weak tests of macroecological theory. *Oikos*, 102(3):679–685, 2003.

[61] Mark M. Meerschaert, David A. Benson, Hans-Peter Scheffler, and Boris Baeumer. Stochastic solution of space-time fractional diffusion equations. *Phys. Rev. E*, 65:041103, Mar 2002.

[62] Mark M. Meerschaert and Farzad Sabzikar. Tempered fractional brownian motion. *Statistics & Probability Letters*, 83(10):2269 – 2275, 2013.

[63] Yasmine Meroz and Igor M. Sokolov. A toolbox for determining subdiffusive mechanisms. *Physics Reports*, 573:1 – 29, 2015. A toolbox for determining subdiffusive mechanisms.

[64] Yasmine Meroz, Igor M Sokolov, and Joseph Klafter. Test for determining a subdiffusive

model in ergodic systems from single trajectories. *Physical review letters*, 110(9):090601, 2013.

[65] Ralf Metzler and Jae-Hyung Jeon. The role of ergodicity in anomalous stochastic processes: analysis of single-particle trajectories. *Physica Scripta*, 86(5):058510, 2012.

[66] Ralf Metzler, Jae-Hyung Jeon, Andrey G Cherstvy, and Eli Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Physical Chemistry Chemical Physics*, 16(44):24128–24164, 2014.

[67] Ralf Metzler and Joseph Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1 – 77, 2000.

[68] Jane Molofsky, Richard Durrett, Jonathan Dushoff, David Griffeath, and Simon Levin. Local frequency dependence and global coexistence. *Theoretical Population Biology*, 55(3):270 – 282, 1999.

[69] A. Mura, M.S. Taqqu, and F. Mainardi. Non-markovian diffusion equations and processes: Analysis and simulations. *Physica A: Statistical Mechanics and its Applications*, 387(21):5033 – 5064, 2008.

[70] Antonio Mura and Francesco Mainardi. A class of self-similar stochastic processes with stationary increments to model anomalous diffusion in physics. *Integral Transforms and Special Functions*, 20(3-4):185–198, 2009.

[71] Antonio Mura and Gianni Pagnini. Characterizations and simulations of a class of stochastic processes to model anomalous diffusion. *Journal of Physics A: Mathematical and Theoretical*, 41(28):285003, 2008.

[72] Gianni Pagnini. The m-wright function as a generalization of the gaussian density for fractional diffusion processes. *Fractional Calculus and Applied Analysis*, 16(2):436–453, 2013.

[73] Gianni Pagnini and Paolo Paradisi. A stochastic solution with gaussian stationary increments of the symmetric space-time fractional diffusion equation. *Fractional Calculus and Applied Analysis*, 19(2):408–440, 2016.

[74] Simone Pigolotti and Massimo Cencini. Speciation-rate dependence in species–area relationships. *Journal of Theoretical Biology*, 260(1):83 – 89, 2009.

[75] Simone Pigolotti and Massimo Cencini. Coexistence and invasibility in a two-species competition model with habitat-preference. *Journal of Theoretical Biology*, 265(4):609 – 617, 2010.

[76] Oscar A Pinto and Miguel A Munoz. Quasi-neutral theory of epidemic outbreaks. *PloS one*, 6(7):e21946, 2011.

[77] FW Preston. Time and space and the variation of species. *Ecology*, 41(4):611–627, 1960.

[78] Robert E. Ricklefs. The unified neutral theory of biodiversity: Do the numbers add up? *Ecology*, 87(6):1424–1431, 2006.

[79] Luca Ridolfi, Paolo D'Odorico, and Francesco Laio. *Noise-induced phenomena in the environmental sciences*. Cambridge University Press, 2011.

[80] Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.

[81] Michael L Rosenzweig et al. *Species diversity in space and time*. Cambridge University Press, 1995.

[82] James Rosindell and Stephen J Cornell. Species-area curves, neutral models, and long-distance dispersal. *Ecology*, 90(7):1743–1750, 2009.

[83] James Rosindell, Stephen J. Cornell, Stephen P. Hubbell, and Rampal S. Etienne. Protracted speciation revitalizes the neutral theory of biodiversity. *Ecology Letters*, 13(6):716–727, 2010.

[84] Hadiseh Safdari, Aleksei V. Chechkin, Gholamreza R. Jafari, and Ralf Metzler. Aging scaled brownian motion. *Phys. Rev. E*, 91:042107, Apr 2015.

[85] Johannes HP Schulz, Eli Barkai, and Ralf Metzler. Aging effects and population splitting in single-particle trajectory averages. *Physical review letters*, 110(2):020602, 2013.

[86] Georg Seisenberger, Martin U Ried, Thomas Endress, Hildegard Büning, Michael Hallek, and Christoph Bräuchle. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science*, 294(5548):1929–1932, 2001.

[87] Yahav Shem-Tov, Matan Danino, and Nadav M Shnerb. Solution of the spatial neutral model yields new bounds on the amazonian species richness. *Scientific reports*, 7:42415, 2017.

[88] Tommaso Spanio, Jorge Hidalgo, and Miguel A. Muñoz. Impact of environmental colored noise in single-species population dynamics. *Phys. Rev. E*, 96:042301, Oct 2017.

[89] Vittoria Sposini, Aleksei V Chechkin, Flavio Seno, Gianni Pagnini, and Ralf Metzler. Random diffusivity from stochastic equations: comparison of two models for brownian yet non-gaussian diffusion. *New Journal of Physics*, 20(4):043044, 2018.

[90] S. Suweis, E. Bertuzzo, L. Mari, I. Rodriguez-Iturbe, A. Maritan, and A. Rinaldo. On species persistence-time distributions. *Journal of Theoretical Biology*, 303:15 – 24, 2012.

[91] Jedrzej Szymanski and Matthias Weiss. Elucidating the origin of anomalous diffusion in crowded fluids. *Physical review letters*, 103(3):038102, 2009.

[92] David Tilman. Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, 101(30):10854–10861, 2004.

[93] Iva Marija Tolić-Nørrelykke, Emilia-Laura Munteanu, Genevieve Thon, Lene Oddershede, and Kirstine Berg-Sørensen. Anomalous diffusion in living yeast cells. *Physical review letters*, 93(7):078102, 2004.

[94] NG Van Kampen. Stochastic processes in chemistry and physics. *Amsterdam: North Holland*, 1:120–127, 1981.

[95] Ilpo Vattulainen and Tomasz Róg. Lipid membranes: Theory and simulations bridged to experiments. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(10):2251 – 2253, 2016. Biosimulations of lipid membranes coupled to experiments.

[96] Aubrey V Weigel, Blair Simon, Michael M Tamkun, and Diego Krapf. Ergodic and nonergodic processes coexist in the plasma membrane as observed by single-molecule tracking. *Proceedings of the National Academy of Sciences*, 108(16):6438–6443, 2011.

[97] Tommaso Zillio, Jayanth R. Banavar, Jessica L. Green, John Harte, and Amos Maritan. Incipient criticality in ecological communities. *Proceedings of the National Academy of Sciences*, 105(48):18714–18717, 2008.

[98] Tommaso Zillio, Igor Volkov, Jayanth R. Banavar, Stephen P. Hubbell, and Amos Maritan. Spatial scaling in model plant communities. *Phys. Rev. Lett.*, 95:098101, Aug 2005.

[99] Robert Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001.

Part II

Second part: Conclusions

Chapter 9

Conclusions

> Avoid doing what you would blame others for doing.
>
> Thales, as quoted in The Lives and Opinions of Eminent Philosophers by Diogenes Laërtius

9.1 First part

For the first part, our first purpose on research was to characterize the process known as ggBm. Firstly, we developed the necessary software to simulate ggBm trajectories and we computed and plotted observables that were important. We were able to obtain also analytical results for all the observables. The first expected result is that it displays anomalous diffusion for its EAMSD as many other stochastic processes and experiments. Then, we calculated the TAMSD and we found two important facts: (i) TAMSD depicts also anomalous diffusion very precisely, and (ii) results from different trajectories were very scattered even for long measurement times. The first fact is an heritage from fBm and the second is due to the random variable in the definition of the ggBm that also appears in the TAMSD. As a consequence of that scattering, the variance of the TAMSD never goes to zero even in the limit of infinite measurement time, and the process is then non-ergodic. We were able to calculate the specific form of the Ergodicity Breaking parameter for any general process that can be written as the product of a random variable distributed according to some distribution multiplied by any ergodic process. According to the Hypothesis section 3.1, we can reformulate the previous result stating that an heterogeneus media that associates a random length scale to each particle can change individual ergodic processes into a non-ergodic one without altering the processes. In particular, we found that for any non-trivial distribution of length scales, there is a continuous transition from ergodicity to non-ergodicity. Nextly, we found that p-variation of ggBm was almost equivalent to that of fBm, except multiplication by a power of the random variable. We obtained that ggBm does not have aging as a consequence of its stationary increments. Then, we decided to study a modified version of the ggBm that featured aging. We characterized the stocastic process and in particular calculated its EAMSD and EATAMSD when some aging time passed between the initiation of the experiment and the start of the measurement. An interesting work would be to find a suitable application for this process now that its main characteristics are known.

9.2 Second part

For the second part, the intention was to study the MSD of several processes mainly by use of the Laplace transforms. Firstly, it was presented the regular overdamped Langevin equation which stochastic force was a power-law correlated fractional Gaussian noise. The result is known to be the fBm. Then, we studied the MSD when this noise was exponentially truncated and, next, power-law truncated as $(1+\tau/\tau_\star)^{-\mu}$. Different behaviours were analyzed depending on the value of μ. At short times, all the models had a MSD that behaves as fBm, ($\sim t^{2H}$). Notably, at long times, the exponentially truncated noise and the strong power-law truncation resulted in normal diffusion. Secondly, we presented the overdamped generalized Langevin equation with fractional Gaussian noise and the Kubo-Zwanzig fluctuations-dissipation relation. After that, we modified the original fractional Gaussian noise with previous exponential truncation and power-law truncation for the autocorrelation function and derived its MSD and velocity autocorrelation function. The short and long time behaviour of the obtained expressions were also calculated. At short times, all models behaved as $\sim t^{2-2H}$. Again, at long times, we obtained normal diffusion for the exponential and strong power-law truncation.

Nextly, instead of tempering the fractional Gaussian noise in the Langevin equation as we previously did, we analysed the model proposed by Meerschaert and Sabzikar in which the tempering is done directly in Mandelbrot's definition of fBm, the so-called tempered fractional Brownian motion. The results presented for this model were published previously by other authors as indicated in the text. At short times, the MSD behaves as fBm. At long times, it converges exponentially to a plateau value. We defined a derivative following the procedure used by Mandelbrot for fractional Gaussian noise that was called tempered fractional Gaussian noise. Then, the autocorrelation function was calculated and we showed that at short times it behaves as fractional Gaussian noise and at long times it decreases exponentially. We also presented the overdamped fractional Langevin equation in which the noise was the tempered fractional Gaussian noise and studied its behaviours. In particular, this motion converges to ballistic diffusion at long times. At the end of this section, we compared the tempered fractional Brownian motion with the Ornstein-Uhlenbeck with fractional Gaussian noise. We showed that they have a similar plateau behaviour at long times and we found the unique value of the Hurst exponent that equals both MSD in the long time limit.

9.3 Third part

Ecosystems have a complex spatial organization that ecologists have tried to describe by observing different patterns of biodiversity across spatial scales. While neutral theory of Ecology can predict some measures in communities of competing species, spatial models still have several problems that limit the quantitative understanding of biodiversity. The purpose on research of this article was to offer the state of the art in spatial models of neutral theory of Ecology. We discussed the main ecological observables and the mathematical predictions of the voter model with speciation. We also performed extensive computational simulations. In particular, we presented a continuous analytical form of the β-diversity in D dimensions, which is the equivalent observable in Ecology to the two-point correlation function in Statistical Physics. We introduced the conditions under which the expression is valid, according to the distances involved and the dispersal kernel of the model. We also discussed the important

case 2D which has special properties not appearing in other dimensions. For the Species-Area Relationships we showed that simulations of the model agree with empirical results. The Species-Abundance Distributions is a hard problem. We presented a standard scaling form valid for any dimension $D \neq 2$ which does not include logarithms and a specific one for the critical dimension $D = 2$ that allows them. The second one was tested with numerical simulations, confirming the need of the use of logarithms to have a better collapse. However, with statistical fluctuations, trivial collapses could be convincing too. Another approach for the Species-Abundance Distributions was presented by introducing a heuristical Fokker-Planck equation for the evolution of the pdf. Logarithmic terms appears and it is suggested that it would be very interesting to compare this approach with the general scaling previously explained. The species persistence times is also discussed and the theoretical behaviours are indicated. In the critical dimension, again, logarithmic terms correct the power-laws. Also, exponential cut-offs appears when the speciation rate is not negligible. These results are corroborated by simulations.

The last part of the article is dedicated to near-neutral models. It is presented a generalization of the voter model, the habitat-preference model which non-neutrality depends on a single parameter. Each node of the lattice have a preference for one type of the species when colonization takes places. It is seen that the non-neutrality leads to an exponentional increase of the average extinction times. This is a stabilizing feature such that species can coexist on realistic time scales. Some generalizations of this model are presented. In particular, it is explained a model in which giving a mild non-neutral preference in only a small fraction of the nodes has as a consequence a durable coexistence. It is discussed how temporally-dependent habitats have a long tradition in Ecology and they have impact on the population growth. Studies show that these models predict better dynamical quantites than neutral models. Finally, models with density dependence are commented.

Part III

Third part: Appendix

Appendix A

Article 1

The reference of this contribution is:

Molina-García, D.; Pham, T. M.; Paradisi, P.; Manzo, C. & Pagnini, G.
Fractional kinetics emerging from ergodicity breaking in random media
Phys. Rev. E, American Physical Society, **2016**, 94, 052147
http://link.aps.org/doi/10.1103/PhysRevE.94.052147

Fractional kinetics emerging from ergodicity breaking in random media

Daniel Molina-García,[1] Tuan Minh Pham,[1,2] Paolo Paradisi,[1,3] Carlo Manzo,[4,*] and Gianni Pagnini[1,5,†]

[1]*BCAM – Basque Center for Applied Mathematics, Alameda de Mazarredo 14, E-48009 Bilbao, Basque Country, Spain*

[2]*Department of Theoretical Physics, Belgorod National Research University, 14 Studencheskaya, 308015 Belgorod, Russia*

[3]*ISTI-CNR, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo," Via Moruzzi 1, I-56124 Pisa, Italy*

[4]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*

[5]*Ikerbasque – Basque Foundation for Science, Calle de Mará Díaz de Haro 3, E-48013 Bilbao, Basque Country, Spain*

(Received 23 July 2015; revised manuscript received 20 October 2016; published 28 November 2016)

We present a modeling approach for diffusion in a complex medium characterized by a random length scale. The resulting stochastic process shows subdiffusion with a behavior in qualitative agreement with single-particle tracking experiments in living cells, such as ergodicity breaking, p variation, and aging. In particular, this approach recapitulates characteristic features previously described in part by the fractional Brownian motion and in part by the continuous-time random walk. Moreover, for a proper distribution of the length scale, a single parameter controls the ergodic-to-nonergodic transition and, remarkably, also drives the transition of the diffusion equation of the process from nonfractional to fractional, thus demonstrating that fractional kinetics emerges from ergodicity breaking.

DOI: 10.1103/PhysRevE.94.052147

I. INTRODUCTION

Many processes in life sciences, soft condensed matter, geology, and ecology show a diffusive behavior that cannot be modeled by classical methods. These phenomena are generally labeled with the term *anomalous* diffusion in order to distinguish them from the *normal* diffusion, where the adjective normal has the double aim of highlighting that (i) a Gaussian-based process is considered, and (ii) that it is a usual diffusion process with a linear growth in time of the particle displacement variance. The observation in nature of anomalous diffusion has been definitively established experimentally and several theoretical models have been proposed for the interpretation of such phenomenon [1–3]. Among these theoretical efforts, the fractional calculus has emerged to be a successful tool for modeling a class of anomalous diffusion processes [4,5]. For this reason, anomalous diffusion governed by equations built on fractional derivatives is often also referred to as *fractional* diffusion. Several stochastic approaches have been proposed in the literature to reproduce fractional kinetics [6–10].

In the last decades, advances in fluorescence-based techniques such as single-particle tracking (SPT) have allowed to precise characterization of the diffusion of molecules in biological systems [11]. In particular, the recording of long single-molecule trajectories has revealed that the occurrence of anomalous diffusion of some cellular components in living cells is associated with ergodicity breaking (EB) [12–16], i.e., the nonequivalence of time and ensemble averages [2,3]. Often EB and anomalous diffusion are concomitant with *aging*, i.e., the dependence of statistical quantities on the measurement time [17].

Besides the fundamental interest of nonergodic processes in statistical mechanics and its still unclear implications in cell biology, the occurrence of EB further embodies a valuable criterium for the selection of the underlying diffusive stochastic process. In this respect, comparative studies involving the fractional Brownian motion (fBm) [18,19], the fractional Langevin equation [18], and the continuous-time random walk (CTRW) [19–21] have been conducted in order to determine which type of motion could possibly cause nonergodic anomalous diffusion. Among the mentioned theoretical frameworks, the fBm and the fractional Langevin motion are ergodic, with fBm displaying EB only in the ballistic limit [18]. On the other hand, the CTRW is nonergodic [20,21], with the EB stemming from the nonstationary nature of the process when the distribution of waiting times has a power-law tail [20]. For this reason, the CTRW has been extensively used to model the occurrence of nonergodic diffusion and the waiting times have been associated to immobilization events caused by biochemical interactions [2,3].

However, due to the lack of nonergodic model alternatives to the CTRW, the use of EB as a criterion to select the dynamic process has shown some limitations. An example is provided by the seminal work of Golding and Cox [12]. In this case, although the presence of EB favors CTRW as the model underlying the dynamics of RNA in cellular cytoplasm, a moments-based criterion called p variation [19] seems to indicate a diffusion compatible with fBm. Similarly, other experiments also showed the simultaneous occurrence of EB and nonlinear scaling of the time-averaged mean-square displacement, making necessary to hypothesize the coexistence of CTRW with other processes in order to theoretically model the observed features [13–15].

In this paper, we provide a general framework in which EB emerges as a consequence of the heterogeneity (or randomness) of the system. The heterogeneity is described by the random nature of a characteristic property of the medium, such as a length scale ℓ_β, depending on a single parameter β. Simple examples of this behavior are provided by a population of particles, each of them diffusing in a Brownian fashion but with a broad distribution of diffusion coefficients ℓ_β. However, our conclusions do not depend on the type of motion performed by the particles. We also show that for any nontrivial choice

*carlo.manzo@icfo.es

†gpagnini@bcamath.org

of the distribution ℓ_β, the parameter β continuously drives the transition from ergodic to a nonergodic process. Notably, a fractional kinetics straightforwardly emerges from EB and thus allows us to associate nonergodicity to a fractional equation.

For its generality, our approach constitutes a flexible tool to interpret the occurrence of EB in random media and in living cells without involving CTRW and subordination. From the biophysical point of view, it implies that EB can be generated by heterogeneity in the diffusion, without the need of particle trapping. In particular, we discuss how our model can resolve the controversy on the interpretation of Golding and Cox experiments [12,19,21,22] by considering the fBm in a heterogeneous medium. Such a model allows one to simultaneously obtain the apparently contradictory features observed in Golding and Cox experiments, i.e., the monotonically increasing p-variation test typical of the fBm together with the EB parameter of the CTRW.

Finally, we show that our formulation can be further generalized by considering a nonstationary length scale $\ell_\beta = \ell_\beta(t)$ and thus including the occurrence of aging.

II. ERGODICITY BREAKING FROM DIFFUSION IN A RANDOM MEDIUM

In our model, we consider a stochastic process defined as

$$X(t) = \ell_\beta \, X_{gen}(t)\,,$$

describing a population of particles diffusing according to a generic ergodic Gaussian process $X_{gen}(t)$ in a complex random medium. The medium properties are independent of the diffusing particles, and its randomness is described by a random characteristic quantity—such as a length scale ℓ_β—with distribution depending on the parameter β. The role of β thus consists in tuning the degree of randomness of the medium by modulating the distribution of the length scale. For the case in which $X_{gen}(t)$ represents a random walk, ℓ_β corresponds to a distribution of diffusion coefficients.

Although the following conclusions hold for every ergodic Gaussian stochastic process, for the sake of simplicity, from now on we will consider the fBm $X_H(t)$, an ergodic non-Markovian Gaussian process characterized by the covariance matrix:

$$\gamma_H(t,s) = t^{2H} + s^{2H} - |t-s|^{2H}\,, \tag{1}$$

where $0 < H < 1$ is the Hurst exponent, and the variance results to be $\langle X_H^2 \rangle = 2\,t^{2H}$.

Therefore, we investigate the following diffusion process $X(t)$ in a random medium:

$$X(t) = \ell_\beta \, X_H(t)\,. \tag{2}$$

In order to study the dynamics of the process, we first consider the time-averaged mean-square displacement [18,20,21]

$$\overline{\delta^2}(T) = \frac{\int_0^{T-\Delta} [X(\xi+\Delta) - X(\xi)]^2 \, d\xi}{T-\Delta}\,, \tag{3}$$

where Δ is the time lag and T the measurement time. The time-averaged mean-square displacement describes the time dependence of the second moment of the particle's position and it is often used to classify the diffusion mode. For the pure Brownian motion ($2H = 1$), $\overline{\delta^2}(T)$ shows a linear growth with

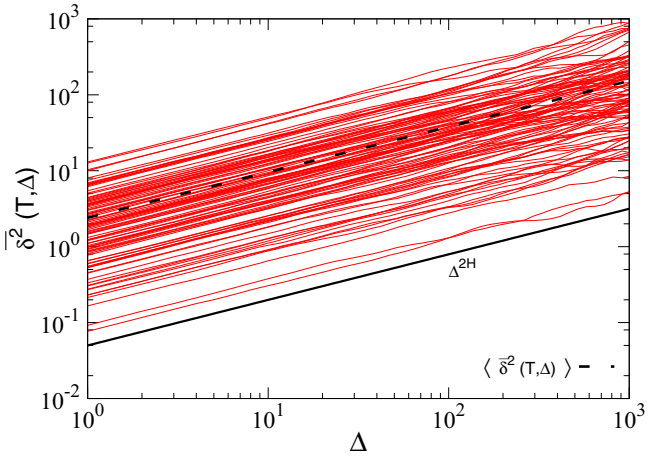


FIG. 1. Time-averaged mean-square displacement $\overline{\delta^2}$ as a function of the time lag Δ calculated for several trajectories (thin red lines) performing the fBm in a random medium, according to the ggBm (7) with $\beta = H = 0.3$ and $T = 10^4$. The dashed line corresponds to the time- and ensemble-averaged mean-square displacement. The continuous thick line is a guide to the eye.

Δ, whereas the fBm shows a power-law behavior $\sim \Delta^{2H}$, i.e., anomalous diffusion. The effect of the random length scale is preserved in the calculation of $\overline{\delta^2}(T)$. For the particular case $2H = 1$ in which the process $X_H(t)$ in (2) corresponds to the pure Brownian motion, the random length scale is proportional to the diffusion coefficient. Consequently, as shown in Fig. 1, time averages such as $\overline{\delta^2}(T)$ remain random variables and is thus irreproducible [23], causing ergodicity breaking. This effect can be estimated through the calculation of the EB parameter $E_B(T)$ [18,21]. Let $\langle\cdot\rangle$ represent the ensemble averaging, then

$$E_B(T) = \frac{\langle [\overline{\delta^2}(T)]^2 \rangle}{\langle \overline{\delta^2}(T) \rangle^2} - 1 \tag{4}$$

is calculated in the large T limit and tends to 0 when the process is ergodic [18].

With a fixed and nonrandom length scale, e.g., $\ell_\beta = 1$, for the stochastic process $X(t)$ defined in (2), we obtain [18]

$$E_B^{(\ell_\beta = 1)}(T) = E_B^{(\mathrm{fBm})}(T) \xrightarrow{T\to\infty} 0\,. \tag{5}$$

In contrast, if ℓ_β is a random variable, for $X(t)$ it holds that

$$E_B^{(\ell_\beta)}(T) = \frac{\langle \ell_\beta^4 \rangle}{\langle \ell_\beta^2 \rangle^2}\left[E_B^{(\mathrm{fBm})}(T) + 1\right] - 1 \xrightarrow{T\to\infty} \frac{\langle \ell_\beta^4 \rangle}{\langle \ell_\beta^2 \rangle^2} - 1\,. \tag{6}$$

The condition $\langle \ell_\beta^4 \rangle > \langle \ell_\beta^2 \rangle^2$ is met in general for any distribution as a consequence of the inequality $K \geqslant S^2 + 1$ [24], where K and S are the kurtosis and the skewness, respectively, and in particular for any unilateral nonincreasing density it holds $K \geqslant 9/5$ from the Gauss-Winckler inequality [24]. The limiting case $\langle \ell_\beta^4 \rangle = \langle \ell_\beta^2 \rangle^2$ is met when the distribution of the length scale is the Bernoulli distribution with equal success probability for values 0 and 1 or it is the Dirac-delta distribution $\delta(\ell_\beta - 1)$; therefore the process is nonergodic for

every nontrivial choice of ℓ_β. Although these conclusions might look somehow trivial, they show how a complex medium, through a random distribution of the length scale, might produce nonergodic behavior into an ergodic Gaussian stochastic process, including the pure Brownian motion, only by introducing heterogeneity [20].

III. ERGODICITY BREAKING AND THE FRACTIONAL KINETICS

In the previous section we have shown that, since ℓ_β is an independent random variable, EB can occur as the sole consequence of the randomness of the medium in which diffusion takes place (6) and independently of the chosen ergodic Gaussian stochastic process.

In the following, we will focus our attention on the stochastic process $X(t)$ as defined in (2). This process has already been studied in a specific characterization named *generalized gray Brownian motion* (ggBm) [8,25,26]. As a matter of fact, the ggBm trajectory $X_{\beta,H}(t)$ is obtained by setting $\ell_\beta = \sqrt{\Lambda_\beta}$, i.e.,

$$X_{\beta,H}(t) = \sqrt{\Lambda_\beta}\, X_H(t)\,, \tag{7}$$

where the positive random variable Λ_β is distributed according to the one-side M-Wright–Mainardi function $M_\beta(\lambda)$, with $\lambda \geqslant 0$ and $0 < \beta < 1$, defined as [27,28]

$$M_\beta(\lambda) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\lambda^k}{\Gamma[-\beta k + (1-\beta)]}\,. \tag{8}$$

The case of a nonrandom length scale, i.e., $\Lambda_\beta = 1$, is straightforwardly recovered in the limit $\beta \to 1$ since it holds $M_1(\lambda) = \delta(\lambda - 1)$. The ggBm is a rather general model and includes as special cases the Brownian motion ($\beta = 2H = 1$), the fBm ($\beta = 1$), and the gray Brownian motion ($\beta = 2H$).

It is well known that the probability density function of $X_{\beta,H}(t)$ is [8]

$$\mathcal{P}(\mathbf{x};\gamma_H) = \frac{1}{\sqrt{(2\pi\lambda)^n \det \gamma_H}} \times \int_0^\infty \exp\left\{-\frac{1}{2\lambda}\,\mathbf{x}^T \gamma_H^{-1} \mathbf{x}\right\} M_\beta(\lambda)\, d\lambda\,, \tag{9}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and $\gamma_H = \gamma_H(t_i, t_j)$, $i, j = 1, \ldots, n$, is the covariance matrix of the fBm defined in (1). Therefore, by the Mellin transform of $M_\beta(\lambda)$ [29], i.e., $\int_0^\infty \lambda^{s-1} M_\beta(\lambda)\, d\lambda = \Gamma[1 + (s-1)]/\,\Gamma[1 + \beta(s-1)]$, with $s > 0$, the covariance matrix of the ggBm can be obtained as [8,26]

$$\gamma_{\beta,H}(t,s) = \frac{1}{\Gamma(1+\beta)}(t^{2H} + s^{2H} - |t-s|^{2H})\,. \tag{10}$$

The one-point one-time density function can be derived from (9) and becomes

$$\mathcal{P}(x;t) = \frac{1}{\sqrt{4\pi\lambda\, t^{2H}}} \int_0^\infty \exp\left\{-\frac{x^2}{4\lambda\, t^{2H}}\right\} M_\beta(\lambda)\, d\lambda \tag{11}$$

$$= \frac{1}{2\,t^H}\, M_{\beta/2}\!\left(\frac{|x|}{t^H}\right), \tag{12}$$

where it emerges that the shape of probability density function of displacements is affected by the medium, here represented by $M_\beta(\lambda)$. In terms of the H function the density function $\mathcal{P}(x;t)$ reads [30,31]

$$\mathcal{P}(x;t) = \frac{1}{2\,t^H}\, H_{01}^{10}\left[\frac{|x|}{t^H}\middle|\begin{matrix} - & ; & (1-\beta/2, \beta/2) \\ (0,1) & ; & - \end{matrix}\right], \tag{13}$$

and the asymptotic decay is $M_{\beta/2}(|x| \to \infty) \sim |x|^{\frac{c}{2}(\beta-1)}\, \mathrm{e}^{-b|x|^c}$, with $b = \frac{2^{1-c}}{c}\, \beta^{\beta c/2}$ and $c = \frac{2}{2-\beta}$ [31,32]. From (10) the variance turns out to be

$$\langle X_{\beta,H}^2 \rangle = \frac{2}{\Gamma(1+\beta)}\, t^{2H}\,, \tag{14}$$

showing that the presence of the medium does not affect the power-law growth of the particle displacement variance over time. It is noteworthy to observe that the ggBm shows both subdiffusion, $0 < H < 1/2$, and superdiffusion, $1/2 < H < 1$. Moreover, a remarkable case is represented by $H = 1/2$ in which the particle displacement variance results to be linear in time, see (14), but the density function is not Gaussian according to (12). The Gaussian density is obtained from (12) as a special case when $\beta = 1$.

The evolution equation for $\mathcal{P}(x;t)$ is given by

$$\frac{\partial \mathcal{P}}{\partial t} = \frac{2H}{\beta} t^{2H-1}\, \mathcal{D}_{2H/\beta}^{\beta-1,1-\beta}\, \frac{\partial^2 \mathcal{P}}{\partial x^2}\,, \tag{15}$$

where $\mathcal{D}_\eta^{\xi,\mu}$ is the Erdélyi-Kober fractional derivative with respect to t and then the process is also referred to as *Erdélyi–Kober fractional diffusion* [33]. Special cases of Eq. (15) are the classical diffusion ($\beta = 2H = 1$), the fBm master equation ($\beta = 1$), and the time-fractional diffusion equation ($\beta = 2H$). A similar approach can be developed in the framework of the space-time–fractional diffusion equation, which includes all its special cases [34].

We would like to remark that the fractional kinetics, i.e., $\beta \neq 1$, emerges directly from the EB due to the randomness of $\ell_\beta = \sqrt{\Lambda_\beta}$ since $M_{\beta\neq 1}(\lambda) \neq \delta(\lambda - 1)$. Moreover, the fractional order related to β can be experimentally computed by means of the long-time limit of the EB parameter. In fact, for large T, from (6) and $\ell_\beta = \sqrt{\Lambda_\beta}$ the EB parameter $E_B^{(\mathrm{ggBm})}(T)$ then becomes

$$E_B^{(\mathrm{ggBm})}(T) \xrightarrow{T\to\infty} \frac{\langle \Lambda_\beta^2 \rangle}{\langle \Lambda_\beta \rangle^2} - 1 = \beta\, \frac{\Gamma(\beta)\Gamma(\beta)}{\Gamma(2\beta)} - 1\,, \tag{16}$$

where again the Mellin transform of $M_\beta(\lambda)$ [29] has been used to compute $\langle \Lambda_\beta^2 \rangle$ and $\langle \Lambda_\beta \rangle$.

In summary, the existence of a random length scale turns an ergodic process into a nonergodic one without the need to introduce an alternative stochastic process. When this transition occurs continuously with respect to a parameter β, the distribution of the length scale can be related to the M-Wright–Mainardi function and the resulting stochastic process is driven by a fractional diffusion equation.

Therefore, the present formulation provides a foundation of fractional kinetics on the basis of the appearance of the EB. In other words, fractional kinetics can be considered as stemming from the EB due to the heterogeneity of the medium in which the diffusion takes place. In order to support this physical foundation argument, we remark that from the proposed ggBm

(7) the evolution of the particle density function is governed by a fractional diffusion equation also in the special case $H = 1/2$, see (15), with $X_H(t)$ performing the classical Brownian motion and the particles displaying a variance with a linear growth in time (14).

IV. RELATION WITH EXPERIMENTS

Advances in biophysical techniques, such as SPT, have allowed researchers to detail the motion of single molecules and have revealed very complex diffusion patterns in living cells [11]. In particular, the analysis of these experiments has shown that several biological systems display nonergodic behavior as a consequence of interactions occurring in heterogeneous cellular environments [12–16]. Such nonergodic behavior has often been connected with the occurrence of anomalous (sub)diffusion. The occurrence of EB has been mainly identified through the nonequivalence of time and ensemble averages and by the calculation of the EB parameter (4) [18,21]. Owing to the importance of molecular transport for the cellular function, theoretical efforts have been devoted to understand the physical mechanism behind EB in biology. Several stochastic models presenting nonstationary (and thus nonergodic) (sub)diffusion have been proposed [3]. Among these models, the most popular has definitively been the CTRW [19–21,35,36], which has been extensively used to model nonergodic subdiffusion in living cells [13–15]. The CTRW has allowed association of the nonergodic behavior with the occurrence of particle immobilization with a heavy-tailed distribution of trapping times [37].

However, among the experimental evidences of EB in biological systems, not all the observed features could be directly addressed within the framework of CTRW alone. For example, Refs. [12,14,15] showed subdiffusive scaling of the time-averaged mean-square displacement obtained for single trajectories, making necessary the postulation of the coexistence of CTRW with other sources of subdiffusion, i.e., the fBm [15] or a fractal processes [14], in order to properly interpret the results. In addition, some experiments did not show the occurrence of inherent features of CTRW, such as *aging* [38] or immobilization [16].

In order to determine the physical scenario behind the subdiffusive EB, a number of diagnostic tools have been proposed [39]. Among these, a valid criterion for selection of stochastic processes is represented by the so-called p-variation test [19]. The test is based on the calculation of the quantity

$$V^{(p)}(t) = \lim_{n\to\infty} V_n^{(p)}(t)\,, \tag{17}$$

where for $t \in [0,T]$,

$$V_n^{(p)}(t) = \sum_{j=0}^{2^n-1} \left| X\left(\frac{(j+1)T}{2^n} \wedge t\right) - X\left(\frac{jT}{2^n} \wedge t\right)\right|^p, \tag{18}$$

with $a \wedge b = \min\{a,b\}$, and allows the CTRW-like models and the fBm to be distinguished, even on the single-trajectory level [39].

In spite of the efforts in developing tests and methods to distinguish between different stochastic models, contradictory indications still prevent the unambiguous determination of the physical mechanism behind EB in biological samples. An example is provided by what is probably the first evidence of EB in living cells, i.e., the experiments describing the motion of individual mRNA molecules inside living *E. coli* cells presented in the seminal paper by Golding and Cox [12]. In this case, in order to explain the occurrence of EB as evidenced by the large scattering of single-trajectory $\overline{\delta^2}$ curves and a nonzero EB parameter, the CTRW was proposed in Refs. [21] and [22] to model this dataset. However, in order to account for the subdiffusive behavior of the time-averaged mean-square displacement, the authors of both works proposed the coexistence of CTRW with some degree of spatial confinement, producing the power-law behavior of $\overline{\delta^2}$ [21,22]. But the application of the p-variation test to the same dataset [12] showed that the subdiffusion is unlikely to originate from the CTRW, whereas the data are compatible with fBm [19].

In this scenario, the general stochastic process presented in this work in (2) provides a plausible framework to describe the subdiffusive nonergodic behavior observed in Ref. [12]. The introduction of a random length scale associate to a random medium allows one to describe the complexity of the cytoskeletal environment and reproduce the scatter of time-average mean-square displacement observed at the single-trajectory level. This observation is quantitatively translated by the calculation of the EB parameter. As a matter of fact, Eq. (16) shows that the EB parameter of the specific process described in (7) is identical to that obtained for a CTRW with a power-law distribution of waiting times, i.e., $\psi(\tau) \propto \tau^{-(1+\beta)}$, and an infinite average sojourn time [21,22], independently of the ergodic Gaussian process used to model diffusion. In addition, the flexibility of our method allows us to choose the fBm to model single-particle diffusion (7) and thus reproduce the subdiffusion in $\overline{\delta^2}$ and maintain the same p-variation behavior of the fBm $V_{ggBm}^{(p)}(t) = \Lambda_\beta^{p/2} V_{fBm}^{(p)}(t)$, while preserving the same degree of EB observed for CTRW-like models (Fig. 2).

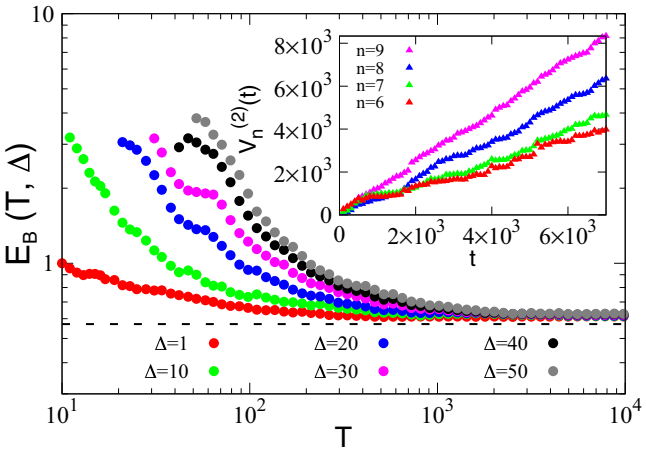


FIG. 2. Plot of $E_B(T)$ for the stochastic process (7) with $\beta = 0.5$ and $H = 0.3$ at various time lags Δ as a function of the measurement time T. Larger Δ produces an increase of $E_B(T)$ at short time T. The $E_B(T)$ values at large time T shown are in agreement with the theoretical expectation (16) (dashed line). (Inset) Results of the the p-variation test with $p = 2$ for the stochastic process (7) with $\beta = 0.5$ and $H = 0.3$, showing the same trend as the pure fBm.

V. AGING

An interesting feature emerging from some single-particle tracking experiments of cellular components [14–16] is the occurrence of *aging*, i.e., the dependence of statistical quantities—such as the time- and ensemble-averaged mean-square displacement $\langle\overline{\delta^2}(T)\rangle$—on the measurement time, as a consequence of the presence of nonstationarity in the diffusive mechanism [17]. Besides living cells, aging has been observed for many complex systems such as blinking nanocrystals [40–42], spin glasses [43], and colloidal suspension [44]. Since aging can characterize long-term memory [45], it can be used as a statistical indicator of complexity and thus exploited to discriminate among different modeling approaches [46,47]. Furthermore, aging has been shown to be associated with weak ergodicity breaking [48,49], i.e., a situation in which the time needed to explore a system phase space is infinite, but the phase space cannot be divided into mutually inaccessible regions [48].

Our theoretical formulation allows reproduction of aging by the extension to the case of a nonstationary random medium $\ell_\beta = \ell_\beta(t)$ [50]. The stochastic process results in the following definition:

$$X_{\alpha,\beta,H}(t) = \sqrt{t^\alpha \Lambda_\beta}\, X_H(t)\,, \tag{19}$$

where Λ_β and $X_H(t)$, with $0 < H < 1$, have the same meaning as in Eq. (7). In this case, the increments of $X_{\alpha,\beta,H}(t)$ are nonstationary, in contrast to the process defined in (7), which is recovered as a particular case for $\alpha = 0$. The parameter α is constrained by the physical requirement that the process is diffusive, meaning that the particle displacement variance must grow in time. Since the variance of the process is given by

$$\left\langle X^2_{\alpha,\beta,H}\right\rangle = \langle\Lambda_\beta\rangle\, t^{\alpha+2H}\,, \tag{20}$$

the latter condition can be expressed as $\alpha > -2H$. It can be shown [50] that the time- and ensemble-averaged mean-square displacement then is $\langle\overline{\delta^2(T)}\rangle \simeq \Delta^{2H} T^\alpha$ (Fig. 3).

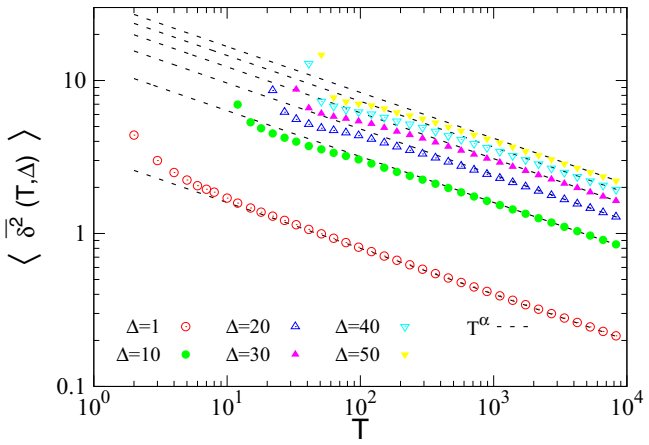


FIG. 3. Plot of the time- and ensemble-averaged mean-square displacement $\langle\overline{\delta^2}(T)\rangle$ at various time lags Δ and as a function of the measurement time T for the process (19) with $\beta = H = 0.3$ and $\alpha = -0.3$. The curves asymptotically show a power-law decay T^α (dashed lines), demonstrating the presence of aging in the process.

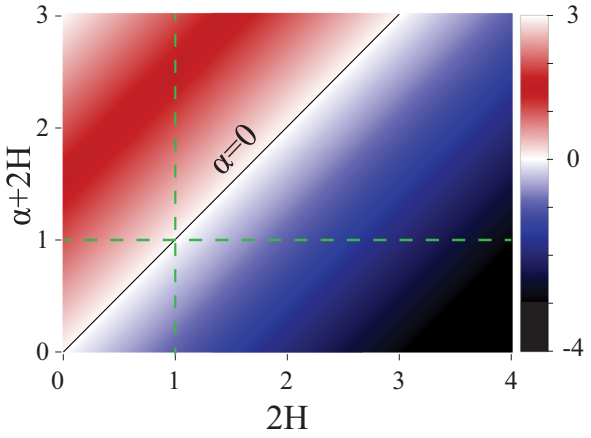


FIG. 4. Contour plot of the aging exponent α as a function of the exponents controlling the power-law growth of the time- ($2H$) and ensemble-averaged ($\alpha + 2H$) mean-square displacement for the process (19). The continuous black line corresponds to the absence of aging ($\alpha = 0$). Dashed green lines separate sub- and superdiffusive regions, characterized by exponent values < 1 and > 1, respectively.

It is interesting to note that our formulation shows properties that were not recapitulated by any of the models for nonergodic diffusion previously presented in literature [3]. First, the exponents controlling the power-law behavior of $\overline{\delta^2}$, $\langle X^2_{\alpha,\beta,H}\rangle$ and $\langle\overline{\delta^2}(T)\rangle$, depend on two parameters, α and H. As such they can thus be independently tuned to reproduce any different scaling of the two curves, in contrast to the other models [3]. In particular, our model show that the time- and ensemble-averaged mean-square displacements can have marked different behavior, for example, with one showing subdiffusivity with the other showing superdiffusivity. In addition, the aging can show a positive or negative exponent depending on the relative magnitude of the exponents controlling the growth of the time- and ensemble-averaged mean-square displacement (Fig. 4).

Moreover, we highlight that the aging can be obtained even in the case in which the time-averaged mean-square displacement $\overline{\delta^2}$ or the ensemble-averaged mean-square displacement $\langle X^2_{\alpha,\beta,H}\rangle$ show Brownian behavior, i.e., when $2H = 1$ or $\alpha = 1$, respectively. It is interesting to note that in the case $2H = 1$ we recover the same relationship between the exponent of the ensemble-averaged mean-square displacement $(\alpha + 1)$ and the time- and ensemble-averaged mean-square displacement obtained for other models, such as the CTRW [20], the scaled Brownian motion [51], the quenched trap [52], and the patch model [53]. Moreover, the calculation of the EB parameter (4) for the process (19) shows that even in the presence of aging ($\alpha \neq 0$) the value of the EB parameter is identical to the one obtained for a CTRW with infinite average sojourn time and power-law distribution of waiting times [50].

VI. CONCLUSIONS

We have demonstrated that an ergodic Gaussian process occurring in a heterogeneous medium characterized by a random length scale can be turned into nonergodic without altering the properties of the Gaussian process itself. We

showed that for any nontrivial choice of the distribution of the length scale, the transition from ergodicity to nonergodicity can be continuously tuned by means of a parameter β. In these cases, the distribution of the length scale can be related to the M-Wright–Mainardi function and the resulting stochastic process is controlled by a fractional diffusion equation.

These conclusions are valid for any ergodic Gaussian process. Therefore, the generality of our formulation posits it as a flexible tool for the interpretation of heterogeneous and/or nonergodic diffusion in disordered systems, such as the many examples of subdiffusion recently observed in living cells [12–16,38]. Notably, our formulation includes the possibility to model the simultaneous occurrence of subdiffusion (as well as any other types of motion) at the single-particle level (Fig. 1) and EB (Fig. 2), a feature observed in many experimental reports [12–15]. This is in contrast with other nonergodic models, such as the CTRW, predicting a linear scaling of the time-averaged mean-square displacement. Therefore, the data could not be satisfactorily interpreted by the CTRW alone and needed to include an additional source of subdiffusion together with CTRW models [14,15].

In particular, we showed that our framework offers an interpretation of the data of Golding and Cox [12] on the basis of a fBm in a heterogeneous medium. The stochastic process (7) allows us to capture both the subdiffusivity in the time-averaged mean-square displacement, the monotonic temporal growing of the p-variations test (as for the fBm), as well as the EB parameter value of the CTRW. Therefore, our model allows us to reproduce all the features observed experimentally and thus solve the disagreement about the underlying stochastic process.

Furthermore, we show that by introducing a nonstationary random medium (19), our model can be extended to include the occurrence of aging, a feature often associated to EB in living systems [14,16]. As such, we consider that our general approach could contribute to investigate the occurrence of EB and anomalous diffusion in life sciences as well as many other fields, and help to elucidate their effects and implications.

ACKNOWLEDGMENTS

This research is supported by the Basque Government through the BERC 2014-2017 program and by the Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa accreditation SEV-2013-0323 and Grant No. MTM2016-76016-R. C.M. acknowledges support from the Spanish Ministry of Economy and Competitiveness ("Severo Ochoa" Programme for Centres of Excellence in R&D (SEV-2015-0522) and FIS2014-56107-R), Fundació Privada CELLEX (Barcelona), HFSP (GA RGP0027/2012), and LaserLab Europe 4 (GA 654148). P.P. acknowledges support from the Bizkaia Talent Organization (2015 Financial Aid Programme for Researchers, Project No. AYD-000-252) and European Commission through the COFUND program.

[1] N. Mercadier, W. Guerin, M. Chevrollier, and R. Kaiser, Nat. Phys. **5**, 602 (2009).
[2] E. Barkai, Y. Garini, and R. Metzler, Phys. Today **65**, 29 (2012).
[3] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, Phys. Chem. Chem. Phys. **16**, 24128 (2014).
[4] D. del Castillo-Negrete, B. A. Carreras, and V. E. Lynch, Phys. Rev. Lett. **94**, 065003 (2005).
[5] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).
[6] D. Fulger, E. Scalas, and G. Germano, Phys. Rev. E **77**, 021122 (2008).
[7] G. Pagnini, Phys. A (Amsterdam, Neth.) **409**, 29 (2014).
[8] A. Mura and G. Pagnini, J. Phys. A: Math. Theor. **41**, 285003 (2008).
[9] R. Gorenflo and F. Mainardi, Eur. Phys. J. Spec. Top. **193**, 119 (2011).
[10] M. Magdziarz, A. Weron, and J. Klafter, Phys. Rev. Lett. **101**, 210601 (2008).
[11] C. Manzo and M. F. Garcia-Parajo, Rep. Prog. Phys. **78**, 124601 (2015).
[12] I. Golding and E. C. Cox, Phys. Rev. Lett. **96**, 098102 (2006).
[13] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unkel, K. Berg-Sørensen, L. Oddershede, and R. Metzler, Phys. Rev. Lett. **106**, 048103 (2011).
[14] A. V. Weigel, B. Simon, M. M. Tamkun, and D. Krapf, Proc. Natl. Acad. Sci. USA **108**, 6438 (2011).
[15] S. M. A. Tabei, S. Burov, H. Y. Kim, A. Kuznetsov, T. Huynh, J. Jureller, L. H. Philipson, A. R. Dinner, and N. F. Scherer, Proc. Natl. Acad. Sci. USA **110**, 4911 (2013).
[16] C. Manzo, J. A. Torreno-Pina, P. Massignan, G. J. Lapeyre, Jr., M. Lewenstein, and M. F. Garcia–Parajo, Phys. Rev. X **5**, 011021 (2015).
[17] S. Burov, R. Metzler, and E. Barkai, Proc. Natl. Acad. Sci. USA **107**, 13228 (2010).
[18] W. Deng and E. Barkai, Phys. Rev. E **79**, 011112 (2009).
[19] M. Magdziarz, A. Weron, K. Burnecki, and J. Klafter, Phys. Rev. Lett. **103**, 180602 (2009).
[20] A. Lubelski, I. M. Sokolov, and J. Klafter, Phys. Rev. Lett. **100**, 250602 (2008).
[21] Y. He, S. Burov, R. Metzler, and E. Barkai, Phys. Rev. Lett. **101**, 058101 (2008).
[22] T. Neusius, I. M. Sokolov, and J. C. Smith, Phys. Rev. E **80**, 011109 (2009).
[23] S. Burov, J.-H. Jeon, R. Metzler, and E. Barkai, Phys. Chem. Chem. Phys. **13**, 1800 (2011).
[24] S. M. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 4th ed., Vol. 1 (C. Griffin & Co., London, 1977).
[25] A. Mura, M. S. Taqqu, and F. Mainardi, Phys. A (Amsterdam, Neth.) **387**, 5033 (2008).
[26] A. Mura and F. Mainardi, Integr. Transf. Spec. F. **20**, 185 (2009).
[27] F. Mainardi, A. Mura, and G. Pagnini, Int. J. Differ. Equations **2010**, 104505 (2010).
[28] G. Pagnini, Fract. Calc. Appl. Anal. **16**, 436 (2013).
[29] F. Mainardi, G. Pagnini, and R. Gorenflo, Fract. Calc. Appl. Anal. **6**, 441 (2003).
[30] F. Mainardi, G. Pagnini, and R. K. Saxena, J. Comp. App. Math. **178**, 321 (2005).

[31] F. Mainardi and G. Pagnini, J. Comp. App. Math. **207**, 245 (2007).
[32] F. Mainardi, Y. Luchko, and G. Pagnini, Fract. Calc. Appl. Anal. **4**, 153 (2001).
[33] G. Pagnini, Fract. Calc. Appl. Anal. **15**, 117 (2012).
[34] G. Pagnini and P. Paradisi, Fract. Calc. Appl. Anal. **19**, 408 (2016).
[35] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
[36] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
[37] G. Bel and E. Barkai, Phys. Rev. Lett. **94**, 240602 (2005).
[38] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini, Phys. Rev. Lett. **103**, 018102 (2009).
[39] Y. Meroz and I. M. Sokolov, Phys. Rep. **573**, 1 (2015).
[40] G. Margolin and E. Barkai, J. Chem. Phys. **121**, 1566 (2004).
[41] P. Paradisi, P. Allegrini, F. Barbi, S. Bianco, and P. Grigolini, AIP Conf. Proc. **800**, 92 (2005).
[42] S. Bianco, P. Grigolini, and P. Paradisi, J. Chem. Phys. **123**, 174704 (2005).
[43] K. Jonason, E. Vincent, J. Hammann, J. P. Bouchaud, and P. Nordblad, Phys. Rev. Lett. **81**, 3243 (1998).
[44] B. Abou, D. Bonn, and J. Meunier, Phys. Rev. E **64**, 021510 (2001).
[45] O. Akin, P. Paradisi, and P. Grigolini, Phys. A (Amsterdam, Neth.) **371**, 157 (2006).
[46] P. Paradisi, R. Cesari, and P. Grigolini, Cent. Eur. J. Phys. **7**, 421 (2009).
[47] O. Akin, P. Paradisi, and P. Grigolini, J. Stat. Mech.: Theory Exp. (2009) P01013.
[48] J. P. Bouchaud, J. Phys. I **2**, 1705 (1992).
[49] J. H. P. Schulz, E. Barkai, and R. Metzler, Phys. Rev. Lett. **110**, 020602 (2013).
[50] D. Molina-García, T. Pham Minh, and G. Pagnini, A stochastic process for the puzzling framework of anomalous diffusion in biological systems, in *Proceedings of the 14th International Conference on Non-Linear Analysis, Non-Linear Systems and Chaos (NOLASC '15), Rome, Italy, November 7–9, 2015*, edited by V. Mladenov, Recent Advances in Electrical Engineering Series Vol. 55 (WSEAS Press, 2015), pp. 75–80.
[51] J.-H. Jeon, A. V. Chechkin, and R. Metzler, Phys. Chem. Chem. Phys. **16**, 15811 (2014).
[52] T. Miyaguchi and T. Akimoto, Phys. Rev. E **83**, 031926 (2011).
[53] P. Massignan, C. Manzo, J. A. Torreno-Pina, M. F. García-Parajo, M. Lewenstein, and G. L. Lapeyre Jr., Phys. Rev. Lett. **112**, 150603 (2014).

Appendix B

Article 2

The reference of this contribution is:

Molina-Garcia, D.; Sandev, T.; Safdari, H.; Pagnini, G.; Chechkin, A. & Metzler, R.
Crossover from anomalous to normal diffusion: truncated power-law noise correlations and applications to dynamics in lipid bilayers
New Journal of Physics, IOP Publishing, **2018**, 20, 103027
https://doi.org/10.1088%2F1367-2630%2Faae4b2

SJR 2017: 1.653
Relative position: #20/202 (Percentile 90%)
Physics and Astronomy: General Physics and Astronomy

Crossover from anomalous to normal diffusion: truncated power-law noise correlations and applications to dynamics in lipid bilayers

Daniel Molina-Garcia[†,♯,$], Trifce Sandev[♭,‡,§], Hadiseh Safdari[£], Gianni Pagnini[†,&], Aleksei Chechkin[♯,¶], and Ralf Metzler[♯]

† BCAM - Basque Center for Applied Mathematics, Alameda de Mazarredo 14, E-48009 Bilbao, Basque Country, Spain

♯ Institute of Physics & Astronomy, University of Potsdam, D-14776 Potsdam-Golm, Germany

$ University of the Basque Country UPV/EHU, Barrio Sarriena s/n, 48940 Leioa, Basque Country, Spain

♭ Radiation Safety Directorate, Partizanski odredi 143, P.O. Box 22, 1020 Skopje, Macedonia

‡ Institute of Physics, Faculty of Natural Sciences and Mathematics, Ss Cyril and Methodius University, P.O. Box 162, 1001 Skopje, Macedonia

§ Research Center for Computer Science and Information Technologies, Macedonian Academy of Sciences and Arts, Bul. Krste Misirkov 2, 1000 Skopje, Macedonia

£ School of Biological Sciences, Institute for Research in Fundamental Sciences, POB 19395-5746, Tehran, Iran

& Ikerbasque – Basque Foundation for Science, Calle de María Díaz de Haro 3, 48013 Bilbao, Basque Country, Spain

¶ Akhiezer Institute for Theoretical Physics, Kharkov 61108, Ukraine

Abstract. The emerging diffusive dynamics in many complex systems shows a characteristic crossover behaviour from anomalous to normal diffusion which is otherwise fitted by two independent power-laws. A prominent example for a subdiffusive-diffusive crossover are viscoelastic systems such as lipid bilayer membranes, while superdiffusive-diffusive crossovers occur in systems of actively moving biological cells. We here consider the general dynamics of a stochastic particle driven by so-called tempered fractional Gaussian noise, that is noise with Gaussian amplitude and power-law correlations, which are cut off at some mesoscopic time scale. Concretely we consider such noise with built-in exponential or power-law tempering, driving an overdamped Langevin equation (fractional Brownian motion) and fractional Langevin equation motion. We derive explicit expressions for the mean squared displacement and correlation functions, including different shapes of the crossover behaviour depending on the concrete tempering, and discuss the physical meaning of the tempering. In the case of power-law tempering we also find a crossover behaviour from faster to slower superdiffusion and slower to faster subdiffusion. As a direct application of our model we demonstrate that the obtained dynamics quantitatively described the subdiffusion-diffusion and subdiffusion-subdiffusion crossover in lipid bilayer systems. We also show that a model of tempered fractional Brownian motion recently proposed by Sabzikar and Meerschaert leads to physically very different behaviour with a seemingly paradoxical ballistic long time scaling.

1. Introduction

Diffusion, the stochastic motion of a tracer particle, was beautifully described by Brown in his study of pollen granules and a multitude of other *molecules* (microscopic particles) [1]. Diffusion is typically described in terms of the mean squared displacement (MSD)

$$\langle x^2(t)\rangle \simeq \mathcal{D}_\alpha t^\alpha \tag{1}$$

of the particle spreading. When $\alpha = 1$ this is the well known law of normal (Brownian or Fickian) diffusion observed in detailed quantitative studies by Perrin, Nordlund, and Kappler [2, 3, 4], among others. In the case of a scaling with an exponent α different from unity, the dynamics encoded by the MSD (1) can be classified in terms of the anomalous diffusion exponent α as either subdiffusive for $0 < \alpha < 1$ or superdiffusive for $\alpha > 1$ [5, 6]. In expression (1) the generalised diffusion coefficient has physical dimension $[\mathcal{D}_\alpha] = \mathrm{cm}^2/\mathrm{s}^\alpha$. Anomalous diffusion with $\alpha \neq 1$ has been revealed in a multitude of systems [5, 6, 7]. In particular, following the massive advances of microscopy techniques anomalous diffusion was discovered in a surging number of biological systems [8, 9]. Thus, subdiffusion was monitored for both endogenous and introduced submicron tracers in biological cells [10, 11, 12, 13, 14, 15, 16, 17] or in inanimate, artificially crowded systems [18, 19, 20]. Supercomputing studies of protein internal motion [21] or of constituent molecules of dilute and protein-crowded lipid bilayer membranes [22, 23, 24, 25, 26] also show subdiffusive behaviour. Due to active motion, also superdiffusion has been reported from several cellular systems [10, 11, 27, 28, 29]. For a more exhaustive list of systems see the recent reviews [8, 9, 30, 31, 32].

In most of these systems the observed anomalous diffusion was identified as fractional Brownian motion or fractional Langevin equation motion type defined below. Both are characterised by power-law correlations of the driving noise [7, 8, 33]. At sufficiently long times, however, this anomalous diffusion will eventually cross over to normal diffusion, when the system's temporal evolution exceeds some relevant correlation time. For instance, all atom molecular dynamics simulations of pure lipid bilayer membranes exhibit a subdiffusive-diffusive crossover at around 10 nsec, the time scale when two lipids mutually exchange their position [22]. The quantitative description of this anomalous-to-normal crossover is the topic of this paper. For both the subdiffusive and superdiffusive situations we include a maximum correlation time of the driving noise and provide exact solutions for the MSD in the case of *hard, exponential* and *power-law* truncation, so-called *tempering*, that can be easily applied in the analysis of experimental or simulations data. The advantage of such a model, in comparison to simply combining an anomalous and a normal diffusive law for the MSD is that the crossover is built into a two-parameter exponential tempering model depending only on the noise strength driving the motion and the crossover time. For the case of a power-law tempering an additional scaling exponent enters. Depending on its magnitude, the anomalous-normal crossover dynamics can be extended to a crossover from either faster to slower superdiffusion or slower to faster subdiffusion. In this approach the crossover between different diffusion regimes thus naturally emerges, and the type of tempering

governs the exact crossover shape. As we will show the crossover shape encoded in this approach nicely fits actual data.

The paper is structured as follows. In section 2 we consider the tempering of superdiffusive fractional Brownian motion and derive the crossover to normal diffusion. In section 3 we perform the same tasks for the subdiffusive generalised Langevin equation. Section 3.5 compares our subdiffusive to normal diffusive model of the tempered generalised Langevin equation to supercomputing data from lipid bilayer membranes exhibiting characteristic crossover dynamics. The data analysis demonstrates excellent agreement with the built-in crossover behaviour of our model. Section 4 addresses direct tempering suggested by Meerschaert and Sabzikar as well as its physicality. Indeed, we show that this type of tempering leads to ballistic motion. We conclude in section 5. Several short appendices provide some additional mathematical details.

2. Tempered superdiffusive fractional Brownian motion

We start from the overdamped stochastic equation of motion of a physical test particle in a viscous medium under the influence of a stochastic force $\xi(t)$ [34, 35]

$$\frac{dx(t)}{dt} = \frac{\xi(t)}{m\eta} = v(t), \tag{2}$$

where $x(t)$ is the particle position and $v(t)$ its velocity. Without loss of generality we assume the initial condition $x(0) = 0$. Furthermore, m is the particle mass, and η, of physical dimension $[\eta] = \mathrm{s}^{-1}$ is the friction coefficient. The stochastic force $\xi(t)$ is assumed to be a stationary and Gaussian noise of zero mean. Then the velocity autocorrelation function fulfils

$$\langle v(t)v(t+\tau)\rangle = \langle v^2\rangle_\tau, \tag{3}$$

for all $\tau \geq 0$. By formal integration of equation (2) the MSD yields in the form

$$\begin{aligned}\langle x^2(t)\rangle &= \int_0^t dt_1 \int_0^t dt_2 \langle v(t_1)v(t_2)\rangle \\ &= 2\int_0^t dt_1 \int_{t_1}^t dt_2 \langle v(t_1)v(t_2)\rangle = 2\int_0^t d\tau (t-\tau)\langle v^2\rangle_\tau.\end{aligned} \tag{4}$$

From this result we infer that if the autocorrelation function $\langle v^2\rangle_\tau$ decays sufficiently fast at long times, such that $\int_0^\infty d\tau \langle v^2\rangle_\tau$ is finite, then the MSD reads

$$\langle x^2(t)\rangle \sim 2t\int_0^\infty d\tau \langle v^2\rangle_\tau, \tag{5}$$

at $t \to \infty$, and diffusion becomes asymptotically normal. Thus, one should expect anomalous diffusion at long times whenever $\int_0^\infty d\tau \langle v^2\rangle_\tau$ is either infinity or zero. This is exactly the case for the persistent and antipersistent fractional Gaussian motions considered in what follows, respectively. In the case of superdiffusive fractional Brownian motion we choose the autocorrelation function in the form

$$\langle v^2\rangle_\tau = \frac{\mathcal{D}_H}{\Gamma(2H-1)}\tau^{2H-2}, \tag{6}$$

where the constant noise strength $\mathcal{D}_H$ has dimension $[\mathcal{D}_H] = \mathrm{cm}^2/\mathrm{s}^{2H}$, $\Gamma(z)$ is the Gamma function, and the Hurst exponent H is in the interval $1/2 \leq H < 1$. We note here that this approach leads to the correct power-law asymptotics of the classical Mandelbrot-van Ness fractional Gaussian noise at long times [36] with $\int_0^\infty d\tau \langle v^2 \rangle_\tau = \infty$, but at the same time leads to an infinite zero-point variance $\langle v^2 \rangle_{\tau=0}$ of the noise.‡ Keeping away from $\tau = 0$ we are allowed to restrict ourselves to the power-law form (6). Furthermore the coefficient $\Gamma(2H-1)$ in equation (6) is introduced to capture the white noise limit. Indeed, due to the property of the δ-function [37]

$$\lim_{H \to 0.5^+} \frac{\tau^{2H-2}}{\Gamma(2H-1)} = \delta(\tau) \tag{7}$$

at $H = 0.5$ and with $\int_0^\infty d\tau \delta(\tau) = 1$ equation (6) reduces to

$$\langle v^2 \rangle_\tau = \mathcal{D} \delta(\tau) \tag{8}$$

with $\mathcal{D}_{1/2} = \mathcal{D}$.§

Now, after plugging result (6) into expression (4) the MSD can be readily calculated, yielding

$$\langle x^2(t) \rangle = \frac{2\mathcal{D}_H}{\Gamma(2H+1)} t^{2H}, \tag{9}$$

which yields sub-ballistic superdiffusion with the anomalous diffusion exponent $\alpha = 2H$, and thus $1 < \alpha < 2$.

In what follows we consider both a hard exponential and a power-law truncation (tempering) of the persistent fractional Gaussian noise with Hurst exponent $1/2 \leq H < 1$.

2.1. Exponentially truncated fractional Gaussian noise

Let us first consider an exponential tempering of the form

$$\langle v^2 \rangle_\tau = \frac{\mathcal{D}_H}{\Gamma(2H-1)} \tau^{2H-2} e^{-\tau/\tau_\star}, \tag{10}$$

for $\tau > 0$, where $\tau_\star > 0$ is a characteristic crossover time scale. For instance, in the case of moving cells the crossover time $\tau_\star$ would correspond to the time scale when the cell motion becomes uncorrelated, similar to the decorrelation of the lipid motion in the example of the lipid bilayer system discussed below.

Here we note that one should keep in mind that the autocorrelation function $\langle v^2 \rangle_\tau$ can not be chosen arbitrary. Namely, its Fourier transform, the spectrum $\langle \tilde{v}^2(\omega) \rangle$ of the

‡ A more consistent approach using the smoothening procedure of fractional Brownian motion over infinitesimally small time intervals à la Mandelbrot and van Ness [36] shows that the weak divergence of the autocorrelation function (6) at $\tau = 0$ does not lead to a change of the MSD.

§ The power-law correlations in the autocorrelation function (6) contrast the sharp δ-correlation of relation (8) [38, 39]. We note that in this combination of the Langevin equation (2) and the autocorrelation function (6) the fluctuation dissipation theorem is not satisfied, and the noise $\xi(t)$ can be considered as an external noise [40], see also the discussion of the generalised Langevin equation below.

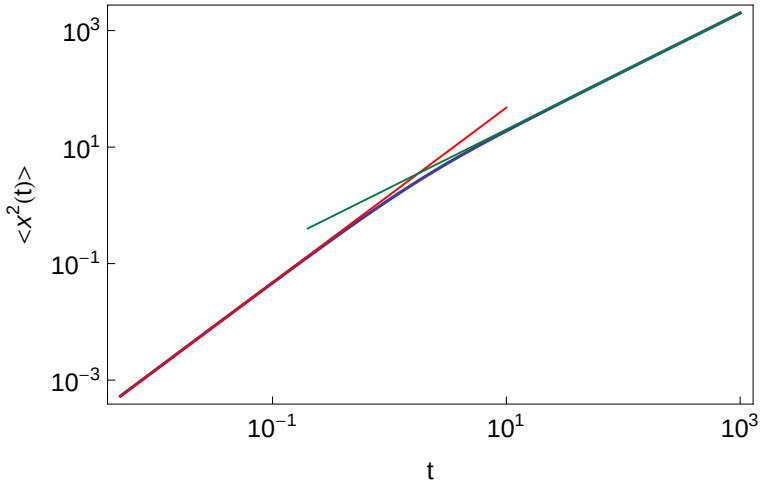


Figure 1. Mean squared displacement (11) for superdiffusive fractional Brownian motion with $H = 3/4$, $\mathcal{D}_H = 1$, and $\tau_\star = 1$ (blue line). The short and long time asymptotics given by expression (12) are depicted by the red and green lines, respectively.

random process $v(t)$ must be non-negative [41]. The positivity of $\langle \tilde{v}^2(\omega) \rangle$ for the case of exponential tempering in equation (10) is shown in Appendix A. Note also that now $\int_0^\infty d\tau \langle v^2 \rangle_\tau = \mathcal{D}_H \tau_\star^{2H-1}$ is finite, thus we expect normal diffusion at long times.

With the use of expression (4) the MSD for the exponentially truncated fractional Gaussian noise takes on the exact form

$$\langle x^2(t) \rangle = \frac{2\mathcal{D}_H \tau_\star^{2H}}{\Gamma(2H-1)} \left[\frac{t}{\tau_\star} \gamma\left(2H-1, \frac{t}{\tau_\star}\right) - \gamma\left(2H, \frac{t}{\tau_\star}\right) \right], \tag{11}$$

where $\gamma(a,z) = \int_0^z t^{a-1} e^{-t} dt$ is the incomplete γ-function. Using the asymptotic $\gamma(a,z) \sim z^a/a$ for $z \ll 1$, and $\gamma(a,z) \sim \Gamma(a)$ for $z \gg 1$, we observe superdiffusive behaviour at short times, and normal diffusion at long times, namely,

$$\langle x^2(t) \rangle \sim \begin{cases} \dfrac{2\mathcal{D}_H}{\Gamma(2H+1)} t^{2H}, & t \ll \tau_\star \\ \\ 2\mathcal{D}_H \tau_\star^{2H-1} t, & t \gg \tau_\star. \end{cases} \tag{12}$$

The emerging normal diffusion thus has the effective diffusivity $\mathcal{D}_H \tau_\star^{2H-1}$. Note that the approximate formula at long times is in concordance with the simple estimate given by expression (5).

Figure 1 shows the crossover behaviour from superdiffusion to normal diffusion encoded in expression (11), along with the short and long time asymptotes given by result (12). As can be discerned from the plot, the crossover region is fairly short, spanning less than a decade in time for the chosen parameters.

2.2. Power-law truncated fractional Gaussian noise

We now consider the softer power-law truncation of the form

$$\langle v^2\rangle_\tau = \frac{\mathcal{D}_H}{\Gamma(2H-1)}\tau^{2H-2}\left(1+\frac{\tau}{\tau_\star}\right)^{-\mu}, \tag{13}$$

for $\tau > 0$, $\mu > 0$ and compare the resulting behaviour with the scenario of exponential tempering. Here, apart from the crossover time $\tau_\star$ the new power-law exponent μ is introduced which effects the dynamics at long times, as we are going to show below. We remark that the positivity of the spectrum for the power-law truncated form is discussed in Appendix A. After plugging (13) into expression (4) we find for the MSD that

$$\langle x^2(t)\rangle = \frac{2\mathcal{D}_H\tau_\star^{2H}}{\Gamma(2H-1)}\left[\frac{t}{\tau_\star}f\left(\mu,2H-1;\frac{t}{\tau_\star}\right)-f\left(\mu,2H;\frac{t}{\tau_\star}\right)\right], \tag{14}$$

where we introduced the notation

$$f(\mu,\alpha;a) = \int_0^a \frac{y^{\alpha-1}}{(1+y)^\mu}dy. \tag{15}$$

Now, using the integral representation [42] of the hypergeometric function ${}_2F_1$ [43] we rewrite the integral in equation (15) as

$$f(\mu,\alpha;a) = \frac{a^\alpha}{\alpha}{}_2F_1(\mu,\alpha,\alpha+1;-a), \tag{16}$$

and thus rewrite the MSD (14) in the final form

$$\langle x^2(t)\rangle = \frac{2\mathcal{D}_H t^{2H}}{\Gamma(2H-1)}\left[\frac{1}{2H-1}{}_2F_1\left(\mu,2H-1;2H;-\frac{t}{\tau_\star}\right)\right.$$
$$\left.-\frac{1}{2H}{}_2F_1\left(\mu,2H;2H+1;-\frac{t}{\tau_\star}\right)\right], \tag{17}$$

In this notation the MSD can be directly evaluated by Wolfram Mathematica [44]. Note that ${}_2F_1(0,b;c;z)=1$, and thus result (17) reduces exactly to the MSD (9) for the untruncated case $\mu = 0$. To obtain the limiting behaviours of the MSD (17) at short times $t \ll \tau_\star$ we use the Gauss hypergeometric series for the function ${}_2F_1$, see 15.1.1 in [42]. As result, to leading order we recover the MSD (9) of untruncated fractional Brownian motion.

At long times $t \gg \tau_\star$ the situation for power-law tempering is actually richer than for the case of exponential tempering. To see this, we first employ the linear transformation formula 15.3.7 in [42] and write expression (17) in the form

$$\langle x^2(t)\rangle = \frac{2D_H\tau_\star^{2H-1}t}{\Gamma(2H-1)}\left[\frac{\Gamma(2H-1)\Gamma(\mu+1-2H)}{\Gamma(\mu)}-\frac{\Gamma(2H+1)\Gamma(\mu-2H)}{2H\Gamma(\mu)}\frac{\tau_\star}{t}\right.$$
$$+\frac{1}{2H-\mu-1}\left(\frac{\tau_\star}{t}\right)^{\mu+1-2H}{}_2F_1\left(\mu,\mu+1-2H;\mu+2-2H;-\frac{\tau_\star}{t}\right)$$
$$\left.-\frac{1}{2H-\mu}\left(\frac{\tau_\star}{t}\right)^{\mu+1-2H}{}_2F_1\left(\mu,\mu-2H;\mu+1-2H;-\frac{\tau_\star}{t}\right)\right]. \tag{18}$$

We consider two possible cases:

2.2.1. Weak power-law truncation, $0 < \mu < 2H - 1 < 1$. In this case the third and fourth terms in the square brackets of expression (18) are dominating and we find

$$\langle x^2(t)\rangle \sim \frac{2D_H\tau_\star^\mu}{(2H-\mu)(2H-1-\mu)\Gamma(2H-1)}t^{2H-\mu} \tag{19}$$

for $t \gg \tau_\star$. Note that in the limit $\mu \to 0$ result (19) reduces to the untruncated formula (9). Thus, since we observe the inequality $2H - \mu > 1$ in the case of weak power-law truncation the dynamics is still superdiffusive, however, with a reduced anomalous diffusion exponent smaller than the value $2H$ in the short time limit.

2.2.2. Strong power-law truncation, $\mu > 2H-1 > 0$. Note that in this case the integral of the velocity autocorrelation function (13) over the whole time domain converges, $\int_0^\infty d\tau\langle v^2\rangle_\tau = \mathcal{D}_H\tau_\star^{2H-1}\Gamma(\mu - 2H + 1)/\Gamma(\mu)$, see 2.2.5.24 in [62]. Thus, with expression (5) we expect a linear time behaviour in the long time limit, whereas the term to next order in (4) gives $\int^t d\tau\tau\langle v^2\rangle_\tau \simeq \int^t d\tau\tau^{2H-1-\mu} \simeq t^{2H-\mu}$, a sublinear contribution since $2H - \mu < 1$. Alternatively, it follows from (18) that the main contribution comes from the first term in the square brackets. Thus, in full accordance with expression (5) we get

$$\langle x^2(t)\rangle \sim \frac{2\mathcal{D}_H\Gamma(\mu - 2H + 1)\tau_\star^{2H-1}t}{\Gamma(\mu)} \tag{20}$$

at $t \gg \tau_\star$.

Finally, for the borderline case $0 < \mu = 2H - 1 < 1$ it is in fact easier to consider equation (17). Making use of formula 7.3.1.81 in [63] we see that the leading contribution comes from the first hypergeometric function in the square brackets in expression (17), as ${}_2F_1(2H-1, 2H-1; 2H; z) \sim \Gamma(2H)\Gamma^{-1}(2H-1)(-z)^{-2H+1}\ln(-z)$. For the MSD we then finally obtain

$$\langle x^2(t)\rangle \sim \frac{2D_H\tau_\star^{2H-1}}{\Gamma(2H-1)}t\ln\left(\frac{t}{\tau_\star}\right). \tag{21}$$

Thus, in this borderline limit between weak truncation (leading to reduced superdiffusion at long times) and strong truncation (normal long time diffusion) we here obtain normal diffusion with a logarithmic correction.

Figure 2 demonstrates that for the power-law tempering the crossover region is significantly enhanced, spanning several orders of magnitude, as compared to the much swifter crossover in the case of exponential tempering.

The MSDs for both cases of exponential and power-law truncation are directly compared in figure 3, along with the time derivative of the MSD. As can be seen, the crossover for the exponential tempering occurs much more rapidly. Thus also the amplitude of the long time Brownian scaling is higher in the case of the power-law tempering for the same value of the crossover time scale $\tau_\star$.

A graphical representation of the correlation functions (6), (10) and (13) is given in figure 4. The exponential cutoff appears more abrupt, as it should. However, this difference will obviously be reduced for larger values of the cutoff exponent μ. To fit

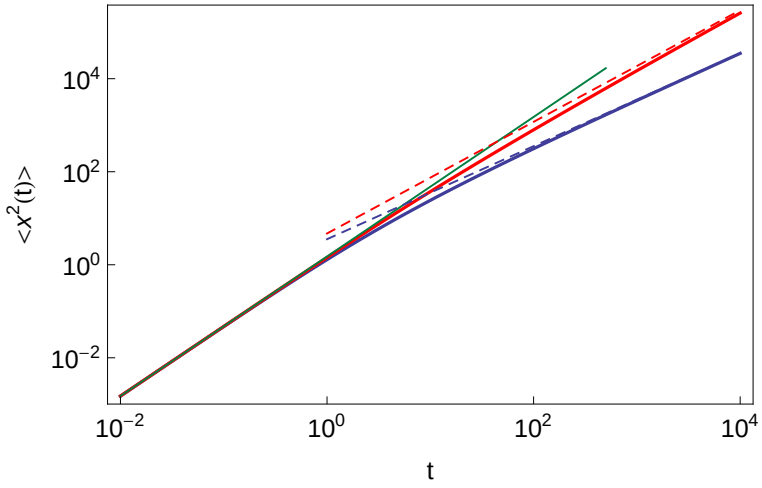


Figure 2. MSD (17) for power-law tempered fractional Brownian motion with $H = 3/4$, $\mathcal{D}_H = 1$, and $\tau_\star = 1$. The red solid line is for $\mu = 0.3$ (weak power law truncation), whereas the blue solid line is for $\mu = 1$ (strong power-law truncation). The red and blue dashed lines correspond to the asymptotics (19) and (20), respectively. The behaviour for the untruncated case given by expression (9) is depicted by the green solid line.

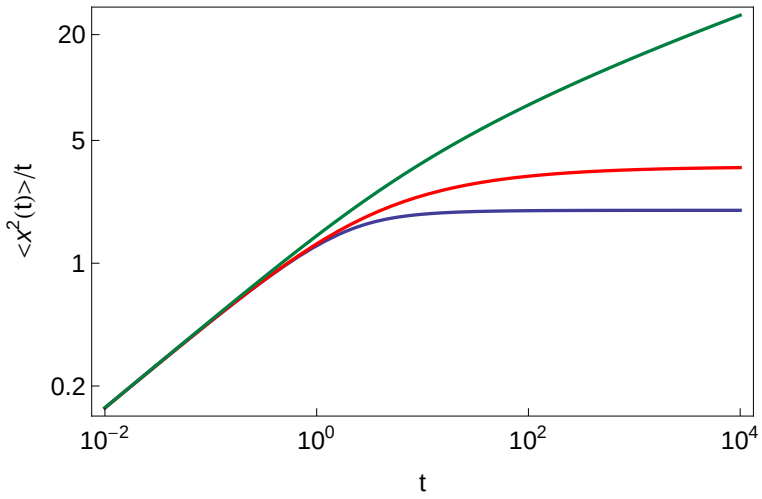


Figure 3. Comparison of the ratio $\langle x^2 \rangle / t$ for different modes of truncation of the power-law noise in equation (2). Parameters: $H = 3/4$, $D_H = 1$, and $\tau_\star = 1$. From bottom to top the blue line depicts the exponential truncation (11) while the red line and green lines show expression (17) for strong ($\mu = 1$) and weak ($\mu = 0.3$) power-law truncation, respectively.

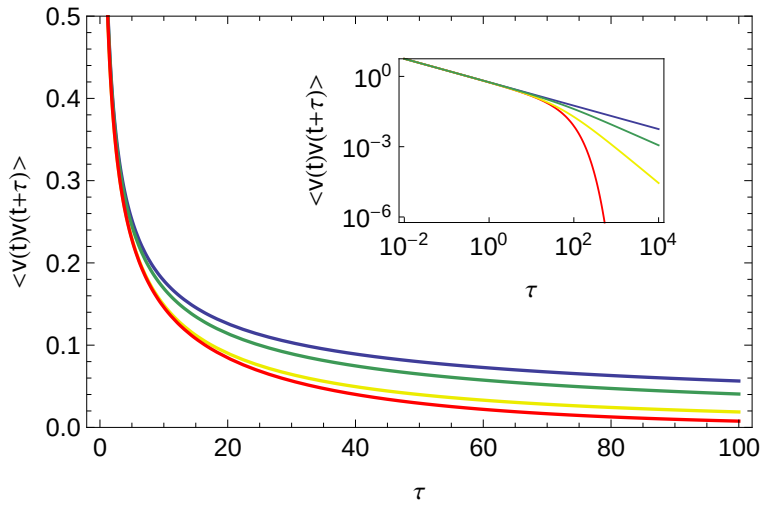


Figure 4. Main figure. Comparison of the velocity autocorrelation functions, from top to bottom: untruncated motion, equation (6) (blue line), weak power-law truncation, equation (13) with $\mu = 0.3$ (green line), strong power-law truncation, equation (13) with $\mu = 1$ (yellow line), and exponential truncation, equation (10) (red line). Parameters: $H = 3/4$, $D_H = 1$, and $\tau_\star = 50$. Inset: double-logarithmic representation.

data, the crossover shape can thus be adjusted by the choice of μ for the case of power-law tempering, thus having the possibility to effect a gradual adjustment from soft power-law to hard exponential tempering.

3. Tempered subdiffusive generalised Langevin equation motion

We now consider the motion encoded in the overdamped generalised Langevin equation for a particle with mass m moving in a viscous medium characterised by the friction kernel $\gamma_H(t)$ of dimension $[\gamma_H(t)] = \mathrm{s}^{-2}$ [7, 38, 45]

$$m \int_0^t \gamma_H(t - t') \frac{dx(t')}{dt'} dt' = \xi(t), \tag{22}$$

where $x(0) = 0$ without loss of generality. Similar to the model considered in section 2 $\xi(t)$ is a Gaussian noise with power-law correlation of the form (6) with $1/2 \leq H < 1$. However, in contrast to the fractional Brownian motion model considered above, we require the system to be thermalised, such that the random force is coupled to the friction kernel through the Kubo-Zwanzig fluctuations dissipation relation [38, 45]

$$\langle \xi^2 \rangle_\tau = k_B T m \gamma_H(\tau). \tag{23}$$

3.1. Mean squared displacement

Let us recall the derivation of the MSD from equations (22) and (23). With our choice $x(0) = 0$ we obtain for the Laplace transform of $x(t)$, $\tilde{x}(s) = \int_0^\infty x(t) \exp(-st) dt$ that

$$\tilde{x}(s) = \frac{\tilde{\xi}(s)}{ms\tilde{\gamma}_H(s)}. \tag{24}$$

Inverse Laplace transformation produces

$$x(t) = \frac{1}{m}\int_0^t \xi(t')H(t-t')dt', \tag{25}$$

where the kernel $H(t)$ is the inverse Laplace transform of $\tilde{H}(s) = 1/[s\tilde{\gamma}_H(s)]$. After some transformation we recover the MSD

$$\begin{aligned}\langle x^2(t)\rangle &= \frac{2}{m^2}\int_0^t dt_1 \int_{t_1}^t dt_2 H(t-t_1)H(t-t_2)\langle \xi^2\rangle_{t_2-t_1} \\ &= \frac{2k_BT}{m}\int_0^t H(t')M(t')dt',\end{aligned} \tag{26}$$

where we introduced $M(t) = \int_0^t \gamma_H(t')H(t-t')dt'$. Its Laplace transform is $\tilde{M}(s) = \tilde{\gamma}_H(s)\tilde{H}(s) = 1/s$, and thus simply $M(t) = 1$. We therefore arrive at

$$\langle x^2(t)\rangle = \frac{2k_BT}{m}\int_0^t H(t')dt'. \tag{27}$$

In Laplace space, this relation reads

$$\tilde{\langle x^2(s)\rangle} = \frac{2k_BT}{m}\frac{\tilde{H}(s)}{s} = \frac{2k_BT}{m}\frac{1}{s^2\tilde{\gamma}_H(s)}. \tag{28}$$

We stop to include a note on when exactly we expect asymptotically normal diffusion in the generalised Langevin equation model. The reasoning is similar to that presented at the beginning of section 2. Namely, from equation (28) it follows that diffusion is normal at long times if $\tilde{\gamma}_H(s)$ tends to a constant in the limit $s \to 0$. This is equivalent to requiring that the average $\int_0^\infty \gamma_H(\tau)d\tau$ is finite or, taking into account the fluctuation-dissipation relation (23) that $\int_0^\infty \langle \xi^2\rangle_\tau d\tau$ is finite (similar to the conclusion in section 2). Then, from expression (28) we infer the following behaviour in the long time limit (compare with equation (5))

$$\langle x^2(t)\rangle = \frac{2k_BT}{m\int_0^\infty \gamma_H(\tau)d\tau}t. \tag{29}$$

According to this, anomalous diffusion is expected at long times whenever $\int_0^\infty \gamma_H(\tau)d\tau$ is either infinite (subdiffusion) or zero (superdiffusion).‖

In accordance with section 2 we choose the friction kernel in the power-law form

$$\gamma_H(\tau) = \frac{\Gamma_H}{\Gamma(2H-1)}\tau^{2H-2}, \tag{30}$$

where the coefficient Γ_H is of dimension $[\Gamma_H] = \mathrm{s}^{-2H}$. The normal Brownian case is recovered from equation (22) for $H = 1/2$ since for $H \to 1/2+$ we see that $\gamma_H(t) \to \Gamma_{1/2}\delta(t)$ (note that in this Brownian limit, $\Gamma_{1/2} = \eta$) and equation (22) assumes the form of the standard Langevin equation driven by white Gaussian noise obeying the regular fluctuation dissipation theorem. We note that the memory kernel

‖ Note here the difference to the results in section 2 where the fluctuation-dissipation theorem is not applied: in that case divergence of the integral over the correlator of the noise $\xi(t)$ over the entire time domain leads to superdiffusion, while subdiffusion emerges when the integral is identical to zero.

for the power-law form (30) can be rewritten in terms of a fractional derivative, and the resulting version of equation (22) is then often referred to as the fractional Langevin equation [7, 46, 47, 48]. Power-law memory kernels of the form (30) are typical for many viscoelastic systems [8, 9, 14, 15, 16, 17, 19, 20, 22, 48].

We now use the Laplace transform of equation (30), $\tilde{\gamma}_H(s) = \Gamma_H s^{1-2H}$, plug this into the above expression, and take an inverse Laplace transformation. This procedure leads to the final result

$$\langle x^2(t)\rangle = \frac{1}{\Gamma(3-2H)}\frac{2k_BT}{m\Gamma_H}t^{2-2H}, \tag{31}$$

which reduces to the classical result $\langle x^2(t)\rangle = 2(k_BT/[m\eta])t$ for normal Brownian motion in the limit $H = 1/2$. Therefore, due to the requirement that the system is thermalised and thus the Kubo-Zwanzig fluctuation theorem is fulfilled, the same noise leads to subdiffusion in this case with anomalous diffusion exponent $\alpha = 2 - 2H$ and $0 < 2 - 2H < 1$. Indeed, due to the coupling in relation (23) large noise values lead to large friction values, and therefore the persistence of the noise is turned into antipersistent diffusion dynamics [7, 46, 48].

3.2. Autocorrelation functions of displacements and velocities

We now derive the autocorrelation function of the displacements, following the procedure laid out by Pottier [49]. First, we note that the double Laplace transform of the correlation function of the random force can be written as

$$\langle\tilde{\xi}(s_1)\tilde{\xi}(s_2)\rangle = k_BTm\int_0^\infty dt_1\int_0^\infty dt_2 e^{-s_1t_1-s_2t_2}\gamma_H(|t_2-t_1|). \tag{32}$$

Then we split the domain of integration over t_2 into the two domains $0 \le t_2 \le t_1$ and $t_1 \le t_2 < \infty$. After introducing $\tau = t_1 - t_2$ and $\tau = t_2 - t_1$ in each domain, respectively, we arrive at

$$\langle\tilde{\xi}(s_1)\tilde{\xi}(s_2)\rangle = k_BTm\frac{\tilde{\gamma}_H(s_1)+\tilde{\gamma}_H(s_2)}{s_1+s_2}. \tag{33}$$

This expression represents the Laplace domain formulation of the fluctuation dissipation theorem (23). By help of equations (33) and (22) we then obtain the double Laplace transform of the displacement correlation function,

$$\langle\tilde{x}(s_1)\tilde{x}(s_2)\rangle = \frac{k_BT}{m}\left(\frac{1/\tilde{\gamma}_H(s_1)}{s_1s_2(s_1+s_2)}+\frac{1/\tilde{\gamma}_H(s_2)}{s_1s_2(s_1+s_2)}\right). \tag{34}$$

In the first term in the parentheses we first take the inverse Laplace transformation over s_2, going from $1/[s_2(s_1+s_2)]$ to $[1-\exp(-s_1t_2)]/s_1$. Exchanging s_2 for s_1 we perform the same operation on the second term. Then we inverse Laplace transform the first term with respect to s_1 and make use of the translation formula $\mathscr{L}_s^{-1}\{\exp(-bs)\mathscr{L}_s\{f(t)\}\} = f(t-b)\Theta(t-b)$, where $b > 0$ and $\Theta(t)$ is the Heaviside step function. As result yields

$$\langle x(t_1)x(t_2)\rangle = \frac{1}{\Gamma(3-2H)}\frac{k_BT}{m\Gamma_H}\Big(t_1^{2-2H}+t_2^{2-2H}-|t_2-t_1|^{2-2H}\Big). \tag{35}$$

The velocity autocorrelation function is obtained by differentiation of this expression,

$$\langle v(t_1)v(t_2)\rangle = \langle v^2\rangle_\tau = -\frac{\sin(\pi[2H-1])\Gamma(2H)}{\pi}\frac{k_BT}{m\Gamma_H}|\tau|^{-2H}, \tag{36}$$

where $\tau = t_2 - t_1$. We see that in the relevant parameter range $1/2 < H < 1$ the velocity autocorrelation is negative, $\langle v^2\rangle_\tau < 0$, reflecting the antipersistent character of the resulting motion.

3.3. Exponentially truncated fractional Gaussian noise

For the exponentially truncated friction kernel and thus noise autocorrelation

$$\gamma_H(\tau) = \frac{\langle \xi^2\rangle_\tau}{k_BTm} = \frac{\Gamma_H}{\Gamma(2H-1)}\tau^{2H-2}e^{-\tau/\tau_\star} \tag{37}$$

we obtain the corresponding Laplace transform

$$\tilde{\gamma}_H(s) = \Gamma_H\left(s+\tau_\star^{-1}\right)^{1-2H}. \tag{38}$$

After plugging this expression into relation (28) and taking the inverse Laplace transformation we obtain

$$\langle x^2(t)\rangle = \frac{2k_BT}{m\Gamma_H}t^{2-2H}E^{1-2H}_{1,3-2H}\left(-\frac{t}{\tau_\star}\right) \tag{39}$$

in terms of the three parameter Mittag-Leffler function $E^\delta_{\alpha,\beta}(z)$ (see Appendix B for its definition and some relevant properties). When the crossover time $\tau_\star$ tends to infinity, $E^\delta_{\alpha,\beta}(0) = 1/\Gamma(\beta)$, and we arrive at result (31) for the untruncated noise. In the limit $H = 1/2$ we have $\delta = 0$ and $E^0_{1,2}(z) = 1/\Gamma(2) = 1$, such that equation (39) reduces to the MSD of normal Brownian motion.

At short times $t \ll \tau_\star$ the MSD (39) reduces to the subdiffusive expression (31), whereas at long times $t \gg \tau_\star$ with the help of $E^{1-2H}_{1,3-2H}(-t/\tau_\star) \sim (t/\tau_\star)^{2H-1}$ (see Appendix A), in accordance with relation (29) the MSD exhibits normal Brownian behaviour,

$$\langle x^2(t)\rangle \sim \frac{2k_BT}{m\Gamma_H\tau_\star^{2H-1}}t. \tag{40}$$

We note that a similar crossover was observed in [50] where a modified three-parameter Mittag-Leffler form for the kernel $\gamma_H(\tau)$ was considered.

The crossover from subdiffusion to normal diffusion in this exponentially tempered generalised Langevin equation picture is shown in figure 5. The crossover behaviour occurs over an interval of the order of a decade in time for the chosen parameters.

Let us now turn to the autocorrelation functions. Using expression (38) in equation (34) we obtain

$$\langle \tilde{x}(s_1)\tilde{x}(s_2)\rangle = \frac{k_BT}{m\Gamma_H}\left(\frac{1}{s_1s_2(s_1+s_2)\left(s_1+\tau_\star^{-1}\right)^{1-2H}} + \frac{1}{s_1s_2(s_1+s_2)\left(s_2+\tau_\star^{-1}\right)^{1-2H}}\right). \tag{41}$$

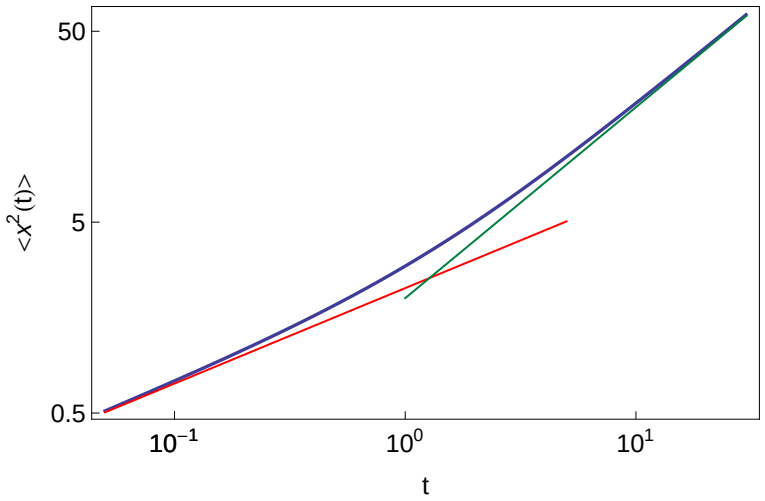


Figure 5. MSD (39) for exponentially tempered generalised Langevin equation motion with $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 1$ (blue line). The short and long time asymptotics (31) and (40) are shown by the red and green lines, respectively.

As above, in the first term in the parentheses we take an inverse Laplace transformation with respect to s_2, and over s_1 in the second term. Then, with the translation formula and the Laplace transform (B.2) of the three parameter Mittag-Leffler function, we find

$$\langle x(t_1)x(t_2)\rangle = \frac{k_BT}{m\Gamma_H}\left(t_1^{2-2H}E^{1-2H}_{1,3-2H}\left(-\frac{t_1}{\tau_\star}\right) + t_2^{2-2H}E^{1-2H}_{1,3-2H}\left(-\frac{t_2}{\tau_\star}\right) \right.$$
$$\left. -|t_2-t_1|^{2-2H}E^{1-2H}_{1,3-2H}\left(-\frac{|t_2-t_1|}{\tau_\star}\right)\right). \tag{42}$$

Differentiation over t_1 and t_2 (with the help of equation (B.6)) then produces the velocity autocorrelation function,

$$\langle v(t_1)v(t_2)\rangle = \langle v^2\rangle_\tau = \frac{k_BT}{m\Gamma_H\tau^{2H}}E^{1-2H}_{1,1-2H}\left(-\frac{\tau}{\tau_\star}\right). \tag{43}$$

with $\tau = t_2 - t_1 > 0$. Using the definition (B.1) of the three parameter Mittag-Leffler function it is easy to check that $E^\delta_{1,\delta}(z) = \exp(z)/\Gamma(\delta)$. Thus, for the velocity autocorrelation function we find the result

$$\langle v^2\rangle_\tau = -\frac{\sin(\pi[2H-1])\Gamma(2H)}{\pi}\frac{k_BT}{m\Gamma_H}\tau^{-2H}e^{-\tau/\tau_\star}, \tag{44}$$

which is anticorrelated and reduces to the untruncated result (36) when the crossover time $\tau_\star$ tends to infinity.

3.4. Power-law truncated fractional noise

For the power-law truncated friction kernel and noise autocorrelator,

$$\gamma_H(\tau) = \frac{\langle\xi^2\rangle_\tau}{k_BTm} = \frac{\Gamma_H}{\Gamma(2H-1)}\tau^{2H-2}\left(1+\frac{\tau}{\tau_\star}\right)^{-\mu} \tag{45}$$

with $\tau > 0$, $\mu > 0$ the Laplace transform of the memory kernel can be performed by use of the integral representation of the Tricomi hypergeometric function $U(a, b; z)$ (see 13.2.5 of [42]), leading to

$$\tilde{\gamma}_H(s) = \Gamma_H \tau_\star^{2H-1} U(2H-1, 2H-\mu; s\tau_\star). \tag{46}$$

With the general relation (28) we thus have

$$\langle x^2(t)\rangle = \frac{2k_BT}{m\Gamma_H \tau_\star^{2H-1}} g(t) \tag{47}$$

with the abbreviation

$$g(t) = \mathscr{L}_s^{-1}\left\{\frac{1}{s^2 U(2H-1, 2H-\mu; s\tau_\star)}\right\}. \tag{48}$$

The inverse Laplace transform of expression (47) cannot be performed analytically. However, we make use of the Tauberian theorems¶ to find the MSD at short and long times.

At short times with $s\tau_\star \gg 1$ we use the large argument asymptotic of the Tricomi function, $U(2H-1, 2H-\mu; s\tau_\star) \sim (s\tau_\star)^{1-2H}$ (13.5.2 in [42]) and thus $\tilde{\gamma}_H(s) \sim \Gamma_H s^{1-2H}$. From equation (28) (or, equivalently, equations (47) and (48)) we then get to result (31) by use of the Tauberian theorem.

Similar to the case considered in section 2 at long times corresponding to $s\tau_\star \ll 1$ the situation is actually richer than for the case of exponential tempering. To see this we first make use of (13.1.3) in [42] to express the Tricomi function via the Kummer function $M(a, b; z)$ through

$$\begin{aligned} U(2H-1, 2H-\mu; s\tau_\star) = \frac{\pi}{\sin(\pi[2H-\mu])}\Bigg[&\frac{M(2H-1, 2H-\mu; s\tau_\star)}{\Gamma(\mu)\Gamma(2H-\mu)} \\ &-(s\tau_\star)^{\mu+1-2H}\frac{M(\mu, \mu+2-2H; s\tau_\star)}{\Gamma(2H-1)\Gamma(\mu+2-2H)}\Bigg]. \end{aligned} \tag{49}$$

Taking into account the series expansion of the Kummer function ((13.1.2) in [42]) we consider the following two possibilities:

3.4.1. Weak power-law truncation, $0 < \mu < 2H-1 < 1$. In this case the second term in (49) is dominant at small s and thus

$$U(2H-1, 2H-\mu; s\tau_\star) \sim \frac{\pi(s\tau_\star)^{1+\mu-2H}}{\sin(\pi[2H-\mu-1])\Gamma(2H-1)\Gamma(\mu+2-2H)}. \tag{50}$$

Plugging this leading behaviour into expressions (47) and (48) and using the Tauberian theorem, after few transformations we obtain the long time behaviour of the MSD,

$$\langle x^2(t)\rangle \sim \frac{\Gamma(2H-1)}{\Gamma(2H-\mu-1)\Gamma(\mu+3-2H)}\frac{2k_BT}{m\Gamma_H\tau_\star^\mu} t^{\mu+2-2H}. \tag{51}$$

¶ The Tauberian theorems state that for slowly varying function $L(t)$ at infinity, i.e. $\lim_{t\to\infty}\frac{L(at)}{L(t)} = 1$, $a > 0$, if $\hat{r}(s) \simeq s^{-\rho}L\left(\frac{1}{s}\right)$, for $s \to 0$, $\rho \geq 0$, then $r(t) = \mathcal{L}^{-1}\left[\hat{r}(s)\right](t) \simeq \frac{1}{\Gamma(\rho)}t^{\rho-1}L(t), \quad t \to \infty$. A similar statement holds for $t \to 0$.

Note that in the limit $\mu \to 0$ expression (51) reduces to the untruncated formula (31). Thus, since we observe the inequality $0 < \mu + 2 - 2H < 1$ in the present case of a weak power-law truncation, the dynamics is still subdiffusive, however, with an anomalous diffusion exponent larger than the value $2 - 2H$ in the short time limit.

3.4.2. Strong power-law truncation, $\mu > 2H - 1 > 0$. In this case the first term in the square brackets in equation (49) becomes dominant at small s and $U(2H-1, 2H-\mu; s\tau_\star) \sim \Gamma(\mu+1-2H)/\Gamma(\mu)$, where we made us of the reflection formula for the Gamma function. From results (47) and (48) by use of the Tauberian theorem we obtain

$$\langle x^2(t)\rangle \sim \frac{\Gamma(\mu)}{\Gamma(\mu+1-2H)} \frac{2k_BT}{m\Gamma_H \tau_\star^{2H-1}} t, \tag{52}$$

valid for $t \gg \tau_\star$. As expected, we find the desired crossover to the normal Brownian scaling of the MSD. Note that this result is in full accordance with equation (29). Indeed, from expression (45) we get (see 2.2.5.24 [62])

$$\int_0^\infty \gamma_H(\tau) d\tau = \frac{\Gamma(\mu+1-2H)}{\Gamma(\mu)} \Gamma_H \tau_\star^{2H-1}. \tag{53}$$

After plugging expression (53) into (29) we arrive at result (52). Note also that the condition of a strong power-law truncation is equivalent to the condition that integral (53) converges.

In the borderline case with $0 < \mu = 2H - 1 < 1$ we use 13.5.9 in [42] and find $U(2H-1, 1; s\tau_\star) \sim -\ln(s\tau_\star)/\Gamma(2H-1)$. With the use of the Tauberian theorem equations (47) and (48) yield

$$\langle x^2(t)\rangle \sim \Gamma(2H-1) \frac{2k_BT}{m\Gamma_H \tau_\star^{2H-1}} \frac{t}{\ln(t/\tau_\star)} \tag{54}$$

at $t \gg \tau_\star$. Thus, in this borderline situation between the cases of weak truncation (leading to increased subdiffusion at long times) and strong truncation (normal long time diffusion) we observe a logarithmic correlation to normal diffusion.

Figure 6 shows the crossover dynamics for power-law tempering for the two possible cases: for weak power-law truncation with $\mu = 0.3$ we observe the predicted crossover from slower to faster subdiffusion, while in the case of strong power-law truncation the subdiffusive dynamics crosses over to normal diffusion.

Figure 7 shows a direct comparison between the cases of exponential and power-law truncation. As expected, the crossover is faster for the exponential tempering, and thus the resulting amplitude in this case exceeds the amplitude for the power-law tempering. Note that the latter observation contrasts the case of the truncated fractional Brownian motion in figure 3, for which the amplitude of the power-law tempering is higher.

3.4.3. Velocity autocorrelation function. To gain some insight into the correlation behaviour we use equation (34) with $\tilde{\gamma}_H(s)$ from equation (46). Taking the inverse

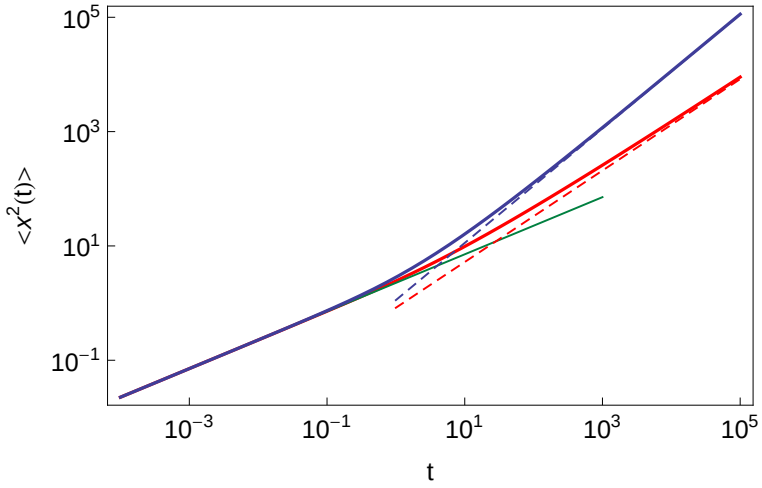


Figure 6. MSD (47) for power-law truncation with $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 1$. The red solid line corresponds to weak power-law truncation with $\mu = 0.3$, the blue solid line to strong truncation with $\mu = 1$. The asymptotics (51) and (52) are shown by red and blue dashed lines, respectively. The thin green solid line corresponds to the MSD (31) for the untruncated case.

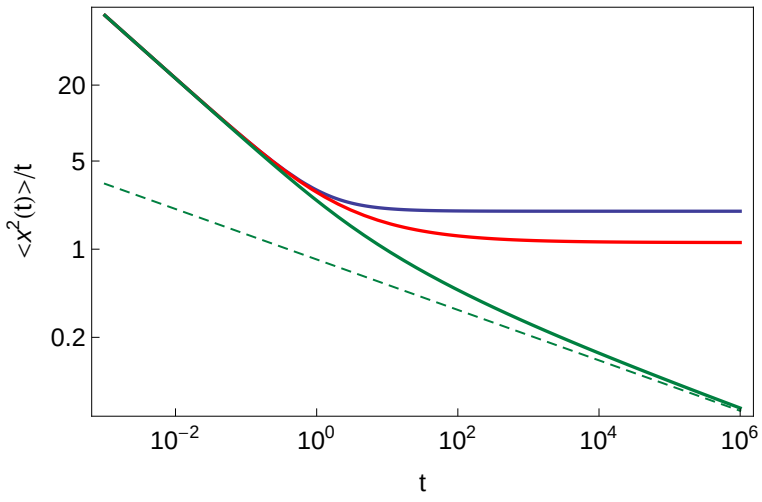


Figure 7. Comparison of the ratio $\langle x^2\rangle/t$ for different truncation modes of the power-law noise in the generalised Langevin equation (22). Parameters: $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 1$. From top to bottom the blue line represents the exponential truncation, equation (39), the red line the strong power-law truncation, equation (47) with $\mu = 1$, and the green line the weak power-law truncation, equation (47) with $\mu = 0.3$. The asymptotics (51) is shown by dashed green line.

Laplace transformation over s_1 and s_2 in the same way as above we obtain the position autocorrelation function

$$\langle x(t_1)x(t_2)\rangle = \frac{k_BT}{m\Gamma_H\tau_\star^{2H-1}}\Big(g(t_1)+g(t_2)-g(|t_2-t_1|)\Big), \tag{55}$$

where $g(t)$ is given by relation (48). From here the velocity autocorrelation function is obtained as

$$\langle v^2\rangle_\tau = \frac{k_BT}{m\Gamma_H\tau_\star^{2H-1}}\frac{d^2}{d\tau^2}g(\tau) \tag{56}$$

with $\tau > 0$. We first note that expression (56) along with (48) may suggest that the Tauberian theorem may be directly applied to the expression $U^{-1}(2H-1, 2H-\mu; s\tau_\star)$ in order to calculate the asymptotic behaviour of the velocity autocorrelation function $\langle v^2\rangle_\tau$. However, for short times corresponding to $s\tau_\star \gg 1$ the function $U^{-1} \sim (s\tau_\star)^{2H-1}$, and since $1/2 < H < 1$, the Tauberian theorem does not apply as $2H-1$ is positive. Instead, we should first obtain the asymptotic of $g(\tau)$ at short times $\tau \ll \tau_\star$ by use of the Tauberian theorem, and only then differentiate twice to get the asymptotic of the velocity autocorrelation function. This way we arrive at expression (36). At long times $\tau \gg \tau_\star$ we again consider the cases of weak and strong power-law truncations separately.

For the weak power-law truncation with $0 < \mu < 2H-1 < 1$ the situation is similar to the short time limit above. Indeed, $U^{-1} \sim (s\tau_\star)^{2H-1-\mu}$, see result (50), and the Tauberian theorem does not apply. Instead we first plug relation (50) into expression (48) and then apply the Tauberian theorem. Following relation (56) we then find

$$\langle v^2\rangle_\tau \sim -C\frac{k_BT}{m\Gamma_H\tau_\star^{\mu}}\frac{1}{\tau^{2H-\mu}}, \tag{57}$$

where $C = (2H-\mu-1)\pi^{-1}\sin(\pi[2H-\mu-1])\Gamma(2H-1)$ is a positive constant. Note that for weak power-law truncation we have $1 < 2H-\mu < 2$, and in the limit $\mu \to 0$ expression (57) reduces to the velocity autocorrelation function (36) in absence of truncation. From comparison of result (57) with (36) we see that the autocorrelation function in the truncated case decays slower than in the untruncated case. This may appear counter-intuitive, however, it is in agreement with the antipersistent character of the fractional Langevin equation model in which the MSD scales like $\simeq t^{2-2H}$ and the velocity autocorrelation function at long times scales as $\simeq -\tau^{-2H}$ for $1/2 < H < 1$. This means that a steeper decay of the velocity autocorrelation function corresponds to a more subdiffusive regime. In other words, when H is closer to $1/2$ (the subdiffusive regime is closer to normal diffusion) then the decay of the autocorrelation function is slower. To see this better consider the effective Hurst index $H_{\rm eff} = H - \mu/2$. Then, for weak power-law truncation the MSD scales like $\simeq t^{2-2H_{\rm eff}}$ with $1/2 < H_{\rm eff} < H < 1$, and the velocity autocorrelation function decays as $\simeq -\tau^- 2H_{\rm eff}$. Thus, in the truncated case the diffusion becomes closer to normal, as it should be, while the velocity autocorrelation function decays slower than in the untruncated case, fully consistent with the antipersistent fractional Langevin equation model.

Now let us turn to the case of strong power-law truncation with $\mu > 2H-1 > 0$ in which for simplicity we assume that $\mu + 1 - 2H \neq n$ where $n \in \mathbb{N}$ is a positive

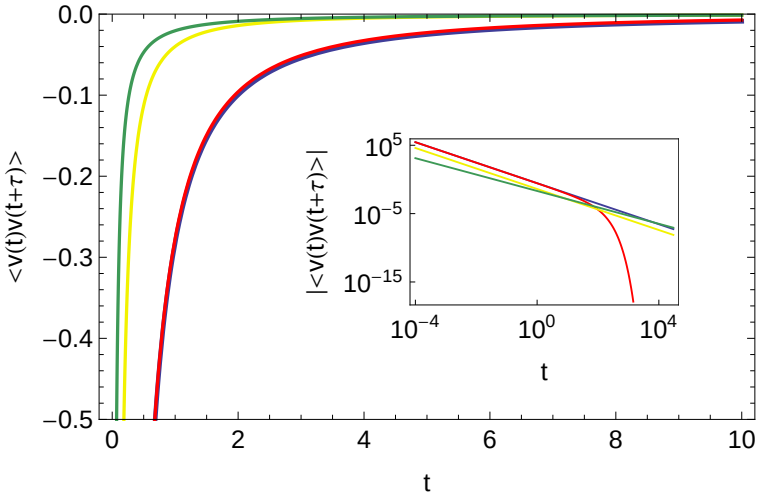


Figure 8. Comparison of the velocity autocorrelation functions for the untruncated case, equation (36) (blue line), with exponential truncation, equation (44) (red line), and with strong power-law truncation, equation (58) where $\mu = 1$ (yellow line), as well as with weak power-law truncation, equation (57) where $\mu = 0.3$ (green line). Parameters: $H = 3/4$, $k_BT/[m\Gamma_H] = 1$, and $\tau_\star = 50$.

integer. We are interested in the exponent of the power-law decay of the velocity autocorrelation function. Then expression (49) yields $U(2H-1, 2H-\mu; s\tau_\star) \sim a_0 + a_1 s + a_2 s^2 + \ldots a_k s^k + a_\mu s^{\mu+1-2H} + a_{k+1} s^{k+1} + \ldots$, where a_i with $i = 0, 1, 2, \ldots$ are constants that can be easily found from expansion 13.1.2 in [42] for the first Kummer function in the square brackets of expression (49) and $k = [\mu + 1 - 2H]$ denotes the integer part of the corresponding argument in the Landau bracket $[\cdot]$. Then $U^{-1}(2H-1, 2H-\mu; s\tau_\star) \sim b_0 + b_1 s + \ldots + b_k s^k + b_\mu s^{\mu+1-2H} + \ldots$ where the b_i with $i = 0, 1, 2, \ldots$ are again constant factors. From here and with equations (48) and (56) we find after application of the Tauberian theorem and subsequent double differentiation

$$\langle v^2 \rangle_\tau \sim -C \frac{k_BT}{m\Gamma_H \tau_\star^{2H-1}} \frac{1}{\tau^{\mu+2-2H}}, \tag{58}$$

where C is a positive constant. Note that in the borderline case $1 > \mu = 2H-1 > 0$ both expressions (57) and (58) tend to the same limit resulting in the logarithmic correction to normal diffusion in expression (54).

A graphical representation of the velocity autocorrelation function (36), (44) and (58) is shown in figure 8.

3.5. Application to lipid molecule dynamics in lipid bilayer membranes

We here demonstrate the usefulness of our tempered fractional Gaussian noise approach to a concrete physical system. The data we have in mind are from all-atom Molecular Dynamics simulations of lipid bilayer membranes [30]. In their simplest form, these are double layered leaves made up of relatively short amphiphilic polymers called lipids. Immersed in water the double layer arrangement prevents the exposure of the

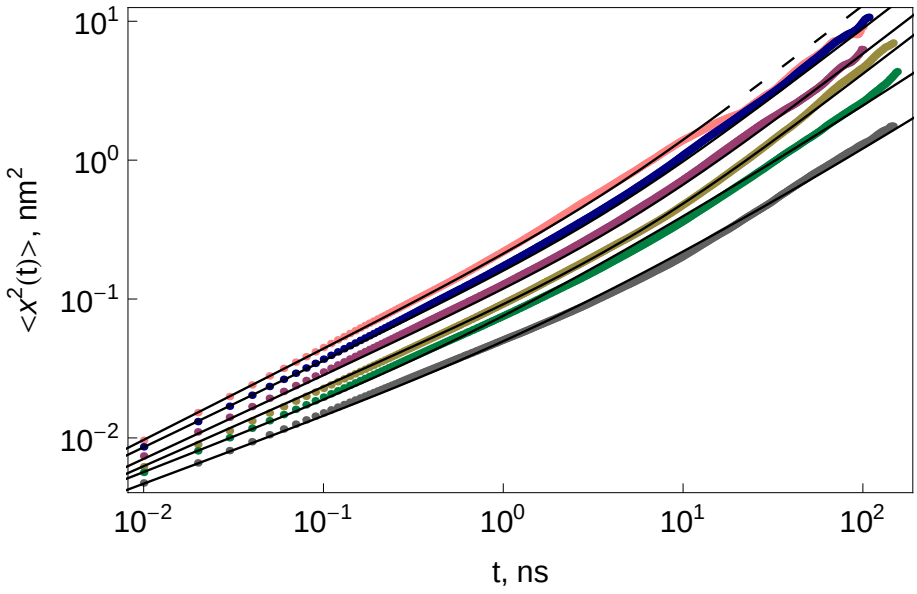


Figure 9. MSD of the motion of lipid molecules in a lipid bilayer model membrane, at room temperature in the liquid disordered and liquid ordered phases (symbols) [22]. The crossover from subdiffusion to normal diffusion or increased subdiffusion at around 10 nsec is distinct. Data courtesy Matti Javanainen, University of Helsinki. The black solid lines provide a fit with equations (39) and (47) resulting from our generalised Langevin equation model with exponentially and power-law truncated noise, respectively. The parameters are presented in Table 1, see also discussion in the text.

hydrophobic tail groups to the ambient water, while the hydrophilic head groups are in contact with the water. At room temperature the lipid bilayer assumes a quite disordered liquid structure [30]. In this lipid matrix, comparatively large membrane proteins may be additionally embedded [30]. Natural biological membranes are composed of lipids of many different chemistries, and they are crowded with membrane proteins. Supercomputing studies have the task to reveal the dynamics of both proteins and lipids in such protein-decorated bilayer systems. This thermally driven diffusion of the constituents influence biological properties of the bilayer, such as diffusion limited aggregation, domain formation, or the membrane penetration by nanoparticles [30].

Figure 9 depicts the simulations results in a chemically uniform, liquid disordered lipid bilayer membrane as well as in the liquid ordered state in the presence of cholesterol molecules—the system is specified in detail in [22]. The motion of the lipids is Gaussian for all cases and best described as viscoelastic diffusion governed by the generalised Langevin equation (22) fuelled by power-law noise [22, 24, 25].[+] As can be seen in figure 9 the MSD of the liquid disordered lipid systems exhibits a clear crossover from subdiffusion to normal diffusion at roughly 10 nsec, the typical crossover time scale

[+] Note that the Gaussian character is lost and intermittent diffusivity dynamics emerge in highly crowded membranes [24], a phenomenon that can be understood in terms of a superstatistical approach [52] or within a fluctuating diffusivity picture [53, 54].

	H	μ	$\tau_\star$ [nsec]	$k_BT/[m\Gamma_H]$ [nm^2/nsec^{2-2H}]	α_{short}	$K_{\alpha_{\text{short}}}$ [nm^2/nsec^{2-2H}]	α_{long}	$K_{\alpha_{\text{long}}}$ [nm^2/nsec$^{\alpha_{\text{long}}}$]
DSPC (purple)	0.70	–	4.0	0.050	0.60	0.034	1.0	0.029
SOPC (pink)	0.67	–	2.5	0.88	0.66	0.064	1.0	0.064
DOPC (blue)	0.69	–	3.0	0.067	0.62	0.046	1.0	0.044
DSPC (grey)	0.76	0.41	0.60	0.019	0.48	0.010	0.89	0.0035
SOPC (green)	0.75	0.44	0.22	0.025	0.50	0.014	0.94	0.0026
DOPC (brown)	0.72	–	4.3	0.038	0.57	0.024	1.0	0.021

Table 1. Fit parameters for the model membrane simulations data shown in figure 9. The colours mentioned in the first column correspond to the colour coding in figure 9.

discussed in literature, at which two nearest neighbour lipid molecules exchange their mutual positions and thus decorrelate their motion [22, 30, 31]. For the liquid ordered cases, one lipid chemistry also shows a subdiffusive-normal crossover, while the two other lipid chemistries lead to a crossover from slower to faster subdiffusion [22]. From fit of the parameters (see the summary in table 1) to the data we observe an excellent agreement with the short and long time scaling regimes and, remarkably, the model fully describes the crossover behaviours without further tuning for both liquid disordered and ordered situations. We note that subdiffusive-diffusive crossovers are also observed for protein-crowded membranes [24, 23, 55].

We note that from equation (31) and the effective diffusion coefficient

$$K^*_\alpha(t) = \frac{1}{2}\frac{d}{dt}\left\langle x^2(t)\right\rangle. \tag{59}$$

we find the short time limiting behaviour

$$K^*_\alpha(t) = K^*_{\alpha,\text{short}} t^{1-2H} \tag{60}$$

with

$$K^*_{\alpha,\text{short}} = \frac{k_BT}{m\Gamma_H}\frac{1}{\Gamma(2-2H)}. \tag{61}$$

For the long time limit, from equation (40), it follows that

$$K^*_{\alpha,\text{long}} = \frac{k_BT}{m\Gamma_H}\frac{1}{\tau_\star^{2H-1}} \tag{62}$$

for the exponential tempering, whereas the cases of DSPC and SOPC lipid chemistries the long time limit in the weak power-law truncation case is given by

$$K^*_{\alpha,\text{long}} = \frac{\Gamma(2H-1)}{\Gamma(2H-\mu+1)\Gamma(\mu+2-2H)}\frac{k_BT}{m\Gamma_H\tau_\star^\mu} \tag{63}$$

The fit values given in table 1 are in very good agreement with those obtained in the simulations study [22]. We note, however, that for the weak power-law tempering model fit the crossover time is somewhat underestimated.

4. Direct tempering of Mandelbrot's fractional Brownian motion

So far we introduced the tempering on the level of the noise $\xi(t)$, which drives the position co-ordinate $x(t)$. Another way to introduce the crossover from anomalous to normal diffusion is to consider a truncation of the power-law correlations directly in the original definition of fractional Brownian motion according to Mandelbrot and van Ness [36]. Such a formulation was recently proposed by Meerschaert and Sabzikar [56]. Here we analyse this model and demonstrate that it leads to a very different behaviour of the MSD than the previous tempered fractional models. A formal mathematical analysis of this model was provided very recently in [57]. We here recall some of their results for the convenience of the reader and present clear physical arguments for the seemingly paradoxical behaviour of this model. In particular we come up with a comparison to a fractional Ornstein-Uhlenbeck scenario.

4.1. Meerschaert and Sabzikar direct tempering model

Meerschaert and Sabzikar defined this extension of fractional Brownian motion by applying an exponential truncating in Mandelbrot's definition [36, 56],*

$$B_{H,\lambda}(t) = \int_{-\infty}^{0} \left[e^{-\lambda(t-t')}(t-t')^{H-\frac{1}{2}} - e^{-\lambda(-t')}(-t')^{H-\frac{1}{2}} \right] B'(t')dt' + \int_{0}^{t} \left[e^{-\lambda(t-t')}(t-t')^{H-\frac{1}{2}} \right] B'(t')dt', \tag{64}$$

where $H, \lambda, t > 0$. $B'(t)$ is white Gaussian noise of δ-covariance $\langle B'(t_1)B'(t_2)\rangle = \sigma^2\delta(t_1 - t_2)$ and zero mean. The parameter λ stands for the truncation parameter, and classical fractional Brownian motion is then obtained in the limiting case $\lambda \to 0$ when $H \in (0, 1)$. It should be noted that the prefactor $1/\Gamma(H + 1/2)$ in Mandelbrot's original definition is dropped here in line with the procedure of [56]. The MSD encoded in equation (64) is (see Appendix C for the derivation)

$$\left\langle B^2_{H,\lambda}(t)\right\rangle = \sigma^2 C_t^2 t^{2H}, \tag{65}$$

where the prefactor is

$$C_t^2 = \left[\frac{2\Gamma(2H)}{(2\lambda t)^{2H}} - \frac{2\Gamma(H+1/2)}{\sqrt{\pi}} \frac{K_H(|\lambda t|)}{(2\lambda t)^H} \right]. \tag{66}$$

$K_H(z)$ denotes the modified Bessel function of the second kind, which for small argument z behaves as [42]

$$K_H(z) \sim \frac{\Gamma(H)}{2^{1-H}} z^{-H} + \frac{\Gamma(-H)}{2^{1+H}} z^H + \frac{\Gamma(H)}{2^{3-H}(1-H)} z^{2-H} \tag{67}$$

while for large z we have $K_H(z) \sim \sqrt{\pi/(2z)}e^{-z}$. The fact that the prefactor C_t^2 is an explicit function of time contrasts the result of standard fractional Brownian motion, and we will readily see the ensuing consequences.

* Note that in this section we use dimensionless units in order not to obfuscate the discussion.

In the short time limit $t \ll \lambda^{-1}$ expression (65) has the compound power-law form

$$\left\langle B^2_{H,\lambda}(t)\right\rangle \sim \sigma^2\Gamma^2(H+1/2)V_H t^{2H} + \frac{\sigma^2\Gamma(2H)}{2^{1+2H}(H-1)}\lambda^{2-2H}t^2 \tag{68}$$

with $V_H = 1/[\Gamma(2H+1)\sin(\pi H)]$. Thus, the limit $\lambda \to 0$ indeed reduces to the expression for standard fractional Brownian motion. In the long time limit $t \gg \lambda^{-1}$ the MSD of this tempered fractional Brownian motion, remarkably, converges exponentially towards a constant value,

$$\left\langle B^2_{H,\lambda}(t)\right\rangle \sim \sigma^2\left(\frac{2\Gamma(2H)}{(2\lambda)^{2H}} - \frac{2^{1/2-H}\Gamma(H+1/2)}{\lambda^{H+1/2}}t^{H-1/2}e^{-\lambda t}\right), \tag{69}$$

a result which is at first surprising. This point will be discussed and compared to the fractional Ornstein-Uhlenbeck process below. The functional behaviour of result (69) is shown in figure 14. We note that if we consider the Langevin equation (2) in combination with the directly tempered noise $B'_{H\lambda}(t)$, expression (65) and its limiting behaviours (68) and (69) exactly correspond to the dynamics of the MSD $\langle x^2(t)\rangle$.

As shown in [57] it is possible to define a tempered fractional Gaussian noise following Mandelbrot and van Ness' smoothening procedure involving a short time lag δ (see Appendix C.2). The autocorrelation function of this tempered fractional Gaussian noise is given through

$$\begin{aligned}\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle = \frac{\Gamma(H+\frac{1}{2})\sigma^2}{\sqrt{\pi}(2\lambda)^H\delta^2}\Big[&2\tau^H K_H(|\lambda\tau|) - (\tau+\delta)^H K_H(\lambda|\tau+\delta|)\\ &- |\tau-\delta|^H K_H(\lambda|\tau-\delta|)\Big].\end{aligned} \tag{70}$$

An important feature of the autocorrelation function (70) for tempered fractional Gaussian noise is its antipersistent behaviour over the whole range $0 < H < 1$ for any finite λ, that is, the integral of expression (70) over the entire domain of τ vanishes:

$$\int_0^\infty \left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle d\tau = 0. \tag{71}$$

This is in sharp contrast to (conventional) fractional Gaussian noise. Indeed, in the limit $\lambda \to 0$ the noise autocorrelation function (70) approaches the one of fractional Gaussian noise [36, 56], as can be derived by using the small argument expansion (67) of the Bessel function. In this limit $\lambda \to 0$ for any finite τ the autocorrelation function (70) converges to

$$\lim_{\lambda\to 0}\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle \sim \frac{\Gamma^2(H+\frac{1}{2})\sigma^2 V_H}{2\delta^2}\left[(\tau+\delta)^{2H} + |\tau-\delta|^{2H} - 2\tau^{2H}\right] \tag{72}$$

and shows negative correlations for $0 < H < 1/2$ and positive correlations for $1/2 < H < 1$, see Appendix C.3.

The autocorrelation function (70) and its limit for $\lambda \to 0$ are shown in figures 10 and 11 for different values of the Hurst parameter. While for the tempered process it is antipersistent for the whole range of H, in the limit $\lambda \to 0$ we clearly see the difference between the antipersistent case with the overshoot to negative values and a slow recovery back to zero. The autocorrelation function for the persistent case is always positive.

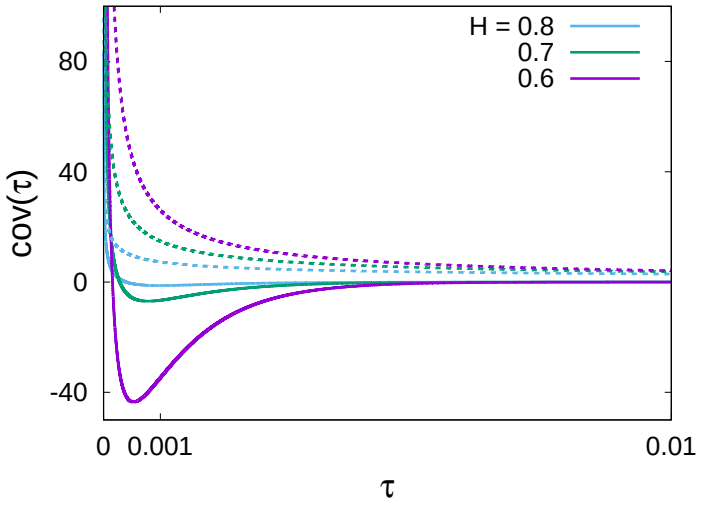


Figure 10. Theoretical results for autocorrelation function, equations (70) and (72), for three different $H > \frac{1}{2}$ values. The solid lines show the antipersistent behaviour of autocorrelation function of tempered fractional Gaussian noise, which approaches zero exponentially; while dashed lines represent the power-law decay of the autocorrelation function of the fractional Gaussian noise. Parameters used: $\lambda = 10^3$, $\delta = 10^{-5}$.

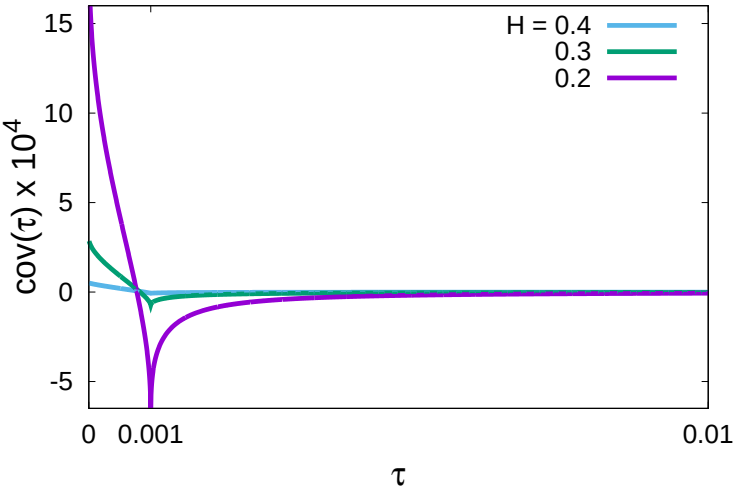


Figure 11. Theoretical results for autocorrelation function, equations (70) and (72), for three different $H < \frac{1}{2}$ values. The solid lines show the autocorrelation function of tempered fractional Gaussian noise and dashed lines are representation of autocorrelation function for fractional Gaussian noise. There is no significant difference between the two functions, except around the truncation time, λ^{-1}, which is magnified in Fig. (12). Parameters used: $\lambda = 10$, $\delta = 10^{-3}$.

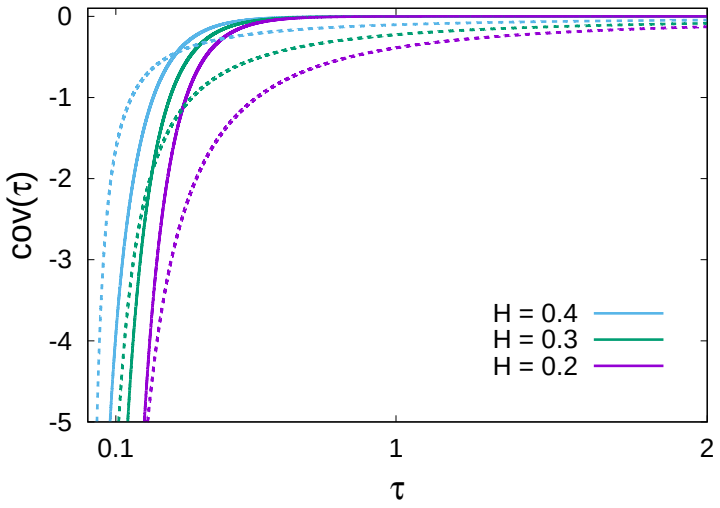


Figure 12. Comparison between the exponentially fast decay of the autocorrelation function of tempered fractional Gaussian noise (solid lines), equation (74), and the slower power-law decay of its ($\lambda \to 0$) regime, which equivalents to fractional Gaussian noise (dashed lines), equation (73), around the truncation time. Parameters used:$\lambda = 10$, $\delta = 10^{-3}$.

It is easy to show that for $\tau \ll 1/\lambda$ and $\delta \to 0$ the autocorrelation function (70) decays as a power law, consistent with the behaviour of fractional Gaussian noise,

$$\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle \sim \sigma^2(2H-1)H\Gamma^2(H+1/2)V_H|\tau|^{2H-2} - \frac{\sigma^2\Gamma(2H)\lambda^{2-2H}}{2^{2H+1}(1-H)}, \tag{73}$$

while the asymptotic behaviour at long observation times, $\tau \gg \lambda^{-1}$,

$$\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle \sim \frac{\tau^{H-\frac{1}{2}}e^{-\lambda\tau}\sigma^2}{2^{H-\frac{1}{2}}\lambda^{H+\frac{1}{2}}\delta^2}\left[1-\cosh(\lambda\delta)+\sinh(\lambda\delta)\frac{(H-\frac{1}{2})\delta}{\tau}\right] \tag{74}$$

decays exponentially, in contrast to the non-tempered limit in equation (73). This different asymptotic behaviour of tempered versus non-tempered fractional Gaussian noise around the truncation time, is shown in figure 12.

4.2. Fractional Langevin equation with directly tempered fractional Gaussian noise

Considering the internal noise $\xi(t)$ of the system as the tempered fractional Gaussian noise $B'_{H,\lambda}(t)$ defined above, the overdamped tempered fractional Langevin equation reads [57]

$$\int_0^t \gamma_H(t-\tau)\frac{dx}{d\tau}d\tau = \xi(t), \tag{75}$$

in which $\gamma_H(\tau) = 2\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle$. Similar to our derivation above, we obtain the Laplace transform of the MSD (28) in dimensionless units,

$$\langle\tilde{x}^2(s)\rangle = \frac{2}{s^2\tilde{\gamma}_H(s)}, \tag{76}$$

in which we have to find the Laplace transformation of the autocorrelation function (70). We assume that $\sigma^2 = 1$ for simplicity from now on. To proceed, in the second and third terms we change the variables and split the resulting integrals,

$$\tilde{\gamma}_H(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H\delta^2}\left\{2[1-\cosh(\delta s)]\int_0^\infty dt e^{-st}t^H K_H(\lambda t)\right. \tag{77}$$

$$\left. + 2\int_0^\delta dt \sinh(s(\delta-t))t^H K_H(\lambda t)\right\}. \tag{78}$$

First, we expand the above functions up to second order in δ. Since in the second integral $\delta \ll \lambda^{-1}$ and $t < \delta$ the relevant regimes are $\delta s \ll 1$ and $\lambda t \ll 1$. Therefore, to second order in δ, $\tilde{\gamma}_H(s)$ is

$$\tilde{\gamma}_H(s) \sim \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H\delta^2}\left\{2\frac{-(\delta s)^2}{2}\int_0^\infty dt e^{-st}t^H K_H(\lambda t)\right.$$

$$\left. +2\int_0^\delta dt(s(\delta-t))t^H K_H(\lambda t)\right\}. \tag{79}$$

Using expansion (67) and keeping terms up to the second order of δ we find

$$2\int_0^\delta dt(s(\delta-t))t^H K_H(\lambda t) \sim \frac{2^H s\pi}{\sin(\pi H)\Gamma(1-H)\lambda^H}\frac{\delta^2}{2}. \tag{80}$$

Insertion of this result back to relation (79) yields

$$\tilde{\gamma}_H(s) \sim \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H\delta^2}\left\{\frac{\pi 2^{H-1}s\delta^2}{\sin(\pi H)\Gamma(1-H)\lambda^H} - (\delta s)^2\int_0^\infty dt e^{-st}t^H K_H(\lambda t)\right\}. \tag{81}$$

The integral in (81) is a Laplace transformation, for which we apply equation (2.16.6.3) of [58]. Hence we find the expression for the autocorrelation function in Laplace space,

$$\tilde{\gamma}_H(s) \sim \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H\delta^2}\left\{\frac{\pi 2^{H-1}s\delta^2}{\sin(\pi H)\Gamma(1-H)\lambda^H}\right.$$

$$\left. -(\delta s)^2\frac{s^{-1}\lambda^{-H}}{2^{H+1}}\sqrt{\pi}\frac{\Gamma(2H+1)}{\Gamma(H+3/2)}\,{}_2F_1\left(\frac{1}{2},1;H+\frac{3}{2};1-\frac{\lambda^2}{s^2}\right)\right\} \tag{82}$$

in terms of the hypergeometric function ${}_2F_1$ [42].

4.2.1. Short time behaviour of the MSD For the regime of short observation times, $\delta \ll t \ll 1/\lambda$ we apply the linear transformations for hypergeometric functions (for more details see (Appendix C.4)). Then, with the general definition for hypergeometric functions up to second order and some simplifications, we find the dominant term for the autocorrelation function,

$$\tilde{\gamma}_H(s) \sim \frac{2\Gamma^2(H+\frac{1}{2})}{2\sin(\pi H)}s^{1-2H}. \tag{83}$$

Substituting this into expression (76), we see that

$$\langle\tilde{x}^2(s)\rangle \sim \frac{2\sin(\pi H)}{2\Gamma^2(H+\frac{1}{2})}s^{2H-3}. \tag{84}$$

By inverse Laplace transformation we find the asymptotic MSD behaviour in time,

$$\langle x^2(t)\rangle \sim \frac{\sin(\pi H)}{\Gamma^2(H+\frac{1}{2})}\frac{t^{2-2H}}{\Gamma(3-2H)}. \tag{85}$$

This result corresponds to subdiffusion for $1/2 < H < 1$ in agreement with the findings in section 3. For $0 < H < 1/2$ the behaviour is superdiffusive.

4.2.2. Long time behaviour of the MSD For the long times regime $t \gg 1/\lambda$ or $\lambda/s \gg 1$ we go back to expression (82) and use the same method as in the previous subsection (see also (Appendix C.5)). It can be seen that the dominant term is a linear function of s,

$$\tilde{\gamma}(s) \sim \frac{\sqrt{\pi}}{2\lambda^{2H}}\frac{2\Gamma(H+\frac{1}{2})}{\sin(\pi H)\Gamma(1-H)}s. \tag{86}$$

Getting back to equation (76) for the MSD, this yields

$$\langle \tilde{x}^2(s)\rangle \sim \frac{2\lambda^{2H}\sin(\pi H)\Gamma(1-H)}{2\Gamma(H+\frac{1}{2})\sqrt{\pi}}s^{-3}. \tag{87}$$

After inverse Laplace transformation, we obtain

$$\langle x^2(t)\rangle \sim \frac{\sin(\pi H)\Gamma(1-H)\lambda^{2H}}{\Gamma(H+\frac{1}{2})\sqrt{\pi}}t^2 = \frac{\sqrt{\pi}\lambda^{2H}}{\Gamma(H+\frac{1}{2})\Gamma(H)}t^2. \tag{88}$$

Thus, at long times this process converges to ballistic diffusion, as already observed in [57].

The general behaviour of the MSD and its crossover from short time power-law behaviour to long time ballistic motion is shown in figure 13 for different Hurst exponents.

4.3. Physical discussion of the direct tempering model and Ornstein-Uhlenbeck with fractional Gaussian noise

To come back to the above observed finite limiting value at long times, encoded in expression (69), of the MSD in the tempered fractional Brownian process we briefly study the confined fractional Brownian motion in an harmonic potential. Experimentally, such a situation arises, for instance, when particle tracking is performed with an optical tweezers setup in a viscoelastic environment [9, 20]. We thus consider the Ornstein-Uhlenbeck process

$$\frac{dx(t)}{dt} = -\lambda x(t) + B'_H(t), \tag{89}$$

for $t > 0$ and with $x(0) = 0$, where the noise $B'_H(t)$ is again fractional Gaussian noise. The MSD reads (see Appendix D)

$$\langle x^2(t)\rangle = \sigma^2 V_H t^{2H} e^{-\lambda t}\left[1 + \frac{\lambda t}{4H+2}\left(e^{\lambda t}f_H(-\lambda t) - e^{-\lambda t}f_H(\lambda t)\right)\right], \tag{90}$$

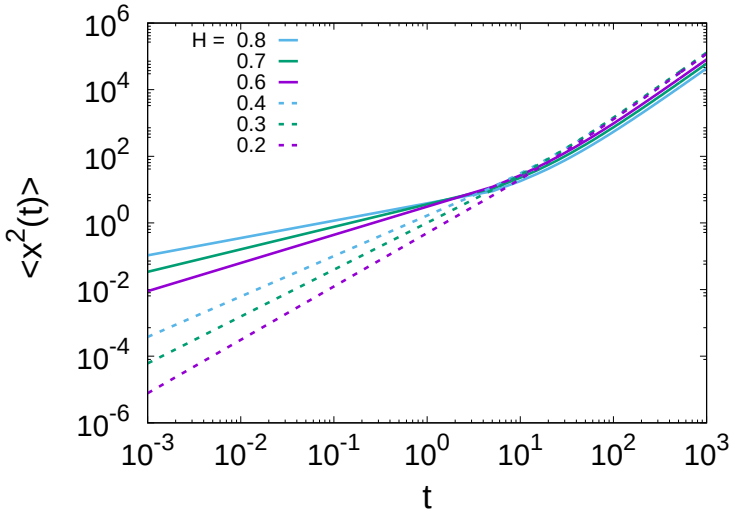


Figure 13. MSD for the tempered Langevin equation (75), from numerical Laplace inversion based on result (82). We also show the transition from anomalous diffusion for short time, equation (85), to the ballistic regime for long observation times, equation (88), is shown for different Hurst exponents and $\lambda = 0.1$.

where $f_H(x) \equiv M(2H+1; 2H+2; x)$ is Kummer's confluent hypergeometric function. For $t \ll \lambda^{-1}$, the MSD of this fractional Ornstein-Uhlenbeck process assumes the form

$$\langle x^2(t)\rangle \sim \sigma^2 V_H t^{2H}(1-\lambda t), \tag{91}$$

which corresponds to unconfined fractional Brownian motion with a correction proportional to λt. In the long-time limit an exponentially fast convergence occurs to the stationary limit

$$\langle x^2(t)\rangle \sim \frac{\sigma^2}{2\sin(\pi H)\lambda^{2H}}. \tag{92}$$

Figure 14 compares the MSDs of tempered fractional Brownian motion and of the fractional Ornstein-Uhlenbeck process. Both of them saturate at long times, where the plateau value depends on the value of H, compare also [59, 60]. Curiously, the plateau values of both processes become identical for the Hurst exponent $H = 0.768149$.

From the comparison with this fractional Ornstein-Uhlenbeck process we see that the direct tempering model of Meerschaert and Sabziker actually describes a confined motion, in contrast to the simple intuition of the tempering in equation (64). In that sense it is fundamentally different from the truncated models considered in the previous sections which show a crossover between two regimes of steadily increasing MSD.

The effect of direct tempering for the fractional Langevin equation model, a priori is even more surprising. Namely, as we saw from equations (85) and (88), this model demonstrates a crossover from a short time subdiffusive to a ballistic regime at long times. Such a behaviour appears counterintuitive. However, as we show not, it is actually a simple consequence of the two basic features of the directly tempered internal

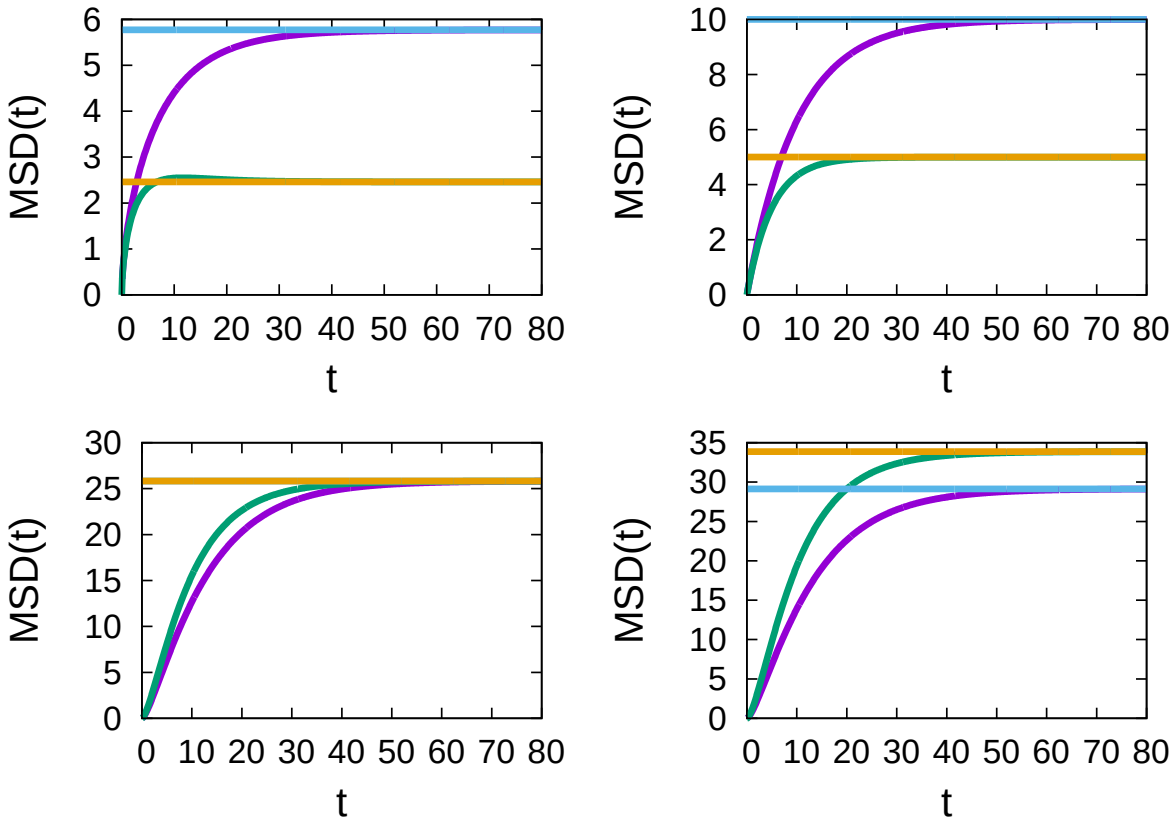


Figure 14. MSD of the tempered fractional Brownian motion (equation (69), violet line) and the fractional Ornstein-Uhlenbeck process (equation (90), green line) and their long time plateaus (horizontal lines) for several values of the Hurst exponent H and the same parameters $\sigma^2 = 1$ and $\lambda = 0.1$. Top left: $H = 0.3$, top right: $H = 0.5$, bottom left: $H = 0.768149$ (equivalence of the plateau values), bottom right: $H = 0.8$.

fractional Gaussian noise (75): (i) the integral of its autocorrelation function over the entire time domain from zero to infinity is identical to zero, see relation (71); (ii) at long times the autocorrelation function exhibits the exponential decay (74). To demonstrate that these two conditions indeed effect the ballistic long time behaviour, consider a toy model for the noise $\xi(t)$ in the fractional Langevin equation (75), namely, we assume the autocorrelation function

$$\langle \xi(t)\xi(t+\tau)\rangle = \gamma_H(\tau) = \delta(\tau) - \lambda e^{-\lambda\tau}. \tag{93}$$

Note that the spectral density of the noise is non-negative and the autocorrelation function (93) obeys conditions (i) and (ii). Now, the Laplace transform of the autocorrelation function (93) reads $\tilde{\gamma}_H(s) = s/(s+\lambda)$, and with relation (76) we thus find the MSD

$$\langle \tilde{x}^2(s)\rangle = \frac{2}{s^2\tilde{\gamma}_H(s)} = \frac{2(s+\lambda)}{s^3}, \tag{94}$$

in Laplace space. As function of time, this indeed produced the ballistic long time behaviour $\langle x^2(t)\rangle \sim \lambda t^2$ for $t \gg 1/\lambda$.

As we see the direct tempering approach leads to unexpected behaviours. Because of the stationary limit (69) the model by Meerschaert and Sabzikar may be more appropriate for modelling the velocity process rather than the position of a diffusing particle. Conversely, the emergence of the ballistic motion (88) at long times for the directly tempered fractional Langevin equation may find useful applications for active systems.

5. Conclusions

In finite systems anomalous diffusion is typically a transient phenomenon, albeit the crossover time to normal diffusive behaviour may be beyond the observation window of the experiment or simulations. In those analyses that explicitly monitor the anomalous-to-normal diffusive crossover, it is desirable to have a complete quantitative model combining the initial anomalous and the terminal normal diffusive regimes, instead of a naive fitting of a non-linear ($\alpha \neq 1$) and a linear ($\alpha = 1$) power-law for the mean squared displacement. The explicit analytical results obtained here provide a two-parameter (exponential cutoff) or three-parameter (power-law cutoff) model for such crossover dynamics and thus have the additional advantage of allowing one to extract the crossover time $\tau_\star$ in those cases when the crossover is rather prolonged and $\tau_\star$ otherwise difficult to extract. Considering systems driven by Gaussian yet power-law correlated noise we introduced two types of tempering of these correlations, a hard exponential and a softer power-law truncation. By plugging this persistent noise into the regular Langevin equation, we produce a superdiffusive-normal diffusive crossover, as would be observed for actively moving but eventually decorrelating particle or animals. In contrast, when we fuel the generalised Langevin equation with this noise, due to the fluctuation dissipation relation the resulting motion becomes antipersistent, and the tempering leads to a subdiffusion-normal diffusion crossover. For the latter case we explicitly showed that the tempered anomalous diffusion model is very useful for the quantitative description of simulations data of lipid molecules in a lipid bilayer membrane. Including the shape of the crossover regime excellent agreement between data and model are observed.

Autocorrelation functions, as studied here, of time series can be directly related to the distribution of first passage times, that is, the distribution of times between consecutive zero crossings of the time series [64]. More recently, the first passage time distribution was studied in the presence of crossovers in the autocorrelation function of the series [65]. In that work the authors demonstrate that the presence of a crossover in the autocorrelation function is related with a crossover in the first passage time distribution which is in fact much more complicated to determine. It will be interesting to explore such a connection for the crossover behaviour studied herein.

We also note here that there exist other classes of anomalous diffusion models such as semi-Markovian continuous time random walks with scale-free waiting time statistic [66], Markovian continuous time random walks with time scale populations

[67], scaled Brownian motion [68], heterogeneous diffusion processes [69], generalised grey Brownian motion [54, 70], or a recent approach using heterogeneous Brownian particle ensembles [71]. The use of either model depends on the physical situation. The motion fuelled by fractional Gaussian noise considered here is useful for a large range of systems, in particular, the motion of submicron tracer particles in living biological cells and artificially crowded environments, or the motion of membrane constituents in pure and protein decorated lipid bilayer membranes. Similarly, applications to stochastic transport in other fields such as sediment transport in earth science [72] are conceivable. To identify such type of motion it is not always sufficient to only look at the MSD of the particle motion, instead, a range of complementary quantitative measures should be considered [7, 32]. To analyse the exact behaviour of these measures for the tempered motion analysed here, including the statistics of time averaged observables [7, 61], will be the focus of future work.

Appendix A. Spectral densities of truncated Gaussian noise

At first we check the positivity of the spectral density of the noise (6). Defining the autocorrelation function $\langle v^2\rangle_\tau$ as symmetric function of the time τ on the infinite axis with respect to $\tau = 0$, the power spectrum becomes

$$\begin{aligned}\langle\tilde{v}^2\rangle_\omega &= \int_{-\infty}^{\infty} d\tau \langle v^2\rangle_\tau e^{\imath\omega\tau} = 2\int_0^\infty d\tau\langle v^2\rangle_\tau \cos(\omega\tau)\\ &= \frac{2\mathcal{D}_H}{\omega^{2H-1}}\sin\left(\left[H-\frac{1}{2}\right]\pi\right),\end{aligned} \tag{A.1}$$

which is positive since $1/2 < H < 1$.

Let us check that for the exponential tempering (10) the spectral density is also positive:

$$\langle\tilde{v}^2\rangle_\omega = \frac{2\mathcal{D}_H}{\left(\omega^2+\tau_\star^{-2}\right)^{H-1/2}}\cos\left([2H-1]\arctan(\omega\tau_\star)\right), \tag{A.2}$$

where we made use of 2.5.31.4 [58]. This expression is non-negative since the argument of the cosine function lies between $-\pi/2$ and $+\pi/2$.

Let us now go to the case of power-law tempering, given by expression (13). Using 2.5.7.6 of [62] we find that

$$\begin{aligned}\langle\tilde{v^2}\rangle_\omega = \frac{2\mathcal{D}_H\tau_\star^{2H-1}}{\Gamma(2H-1)}\Bigg[&\frac{\Gamma(2H-1)\Gamma(\mu-2H+1)\,{}_2F_3\left(\frac{2H-1}{2},H;\frac{1}{2},\frac{2H-\mu}{2},\frac{2H-\mu+1}{2};-\frac{(\omega\tau_\star)^2}{4}\right)}{\Gamma(\mu)}\\ &+\frac{\Gamma(2H-\mu-1)}{(\omega\tau_\star)^{2H-\mu-1}}\cos\left(\frac{[2H-\mu-1]\pi}{2}\right)\\ &\times{}_2F_3\left(\frac{\mu}{2},\frac{\mu+1}{2};\frac{1}{2},\frac{\mu-2H+3}{2},\frac{\mu-2H+2}{2};-\frac{(\omega\tau_\star)^2}{4}\right)\\ &+\mu\,\Gamma(2H-\mu-2)\sin\left(\frac{[\mu-2H+1]\pi}{2}\right)\end{aligned}$$

$$\times \frac{{}_2F_3\left(\frac{\mu+1}{2}, \frac{\mu+2}{2}; \frac{3}{2}, \frac{\mu-2H+4}{2}, \frac{\mu-2H+3}{2}; -\frac{(\omega\tau_\star)^2}{4}\right)}{(\omega\tau_\star)^{2H-\mu-2}}\Bigg]. \tag{A.3}$$

The positivity of this expression was checked numerically with Mathematica for various values of the exponent μ.

We note that since ${}_pF_q\left((a_p);(b_q);0\right) = 1$ [43], we have

$$\lim_{\omega\to 0}\langle\tilde{v}^2\rangle_\omega = 2\Gamma(2-2H)\mathcal{D}_H\tau_\star^{2H-1} > 0 \tag{A.4}$$

for all μ. Moreover, for $\mu = 1$ result (A.3) can be simplified with the use of the following property of the generalised hypergeometric function ([63] 7.2.3.7): if for r values of a_p there also exist equal them r values of b_q, then

$${}_pF_q\left((a_{p-r}),(c_r);(b_{q-r}),(c_r);z\right) = {}_{p-r}F_{q-r}\left((a_{p-r});(b_{q-r});z\right). \tag{A.5}$$

Appendix B. Mittag-Leffler functions and derivation of equation (39)

The three parameter Mittag-Leffler function is defined by [73]

$$E^\delta_{\alpha,\beta}(z) = \sum_{k=0}^{\infty}\frac{(\delta)_k}{\Gamma(\alpha k+\beta)}\frac{z^k}{k!}, \tag{B.1}$$

where $(\delta)_k = \Gamma(\delta+k)/\Gamma(\delta)$ is the Pochhammer symbol. Its Laplace transform is given by [73]

$$\mathcal{L}\left[t^{\beta-1}E^\delta_{\alpha,\beta}(-\nu t^\alpha)\right](s) = \frac{s^{\alpha\delta-\beta}}{\left(s^\alpha+\nu\right)^\delta}, \tag{B.2}$$

where $\mathrm{Re}(s) > |\nu|^{1/\alpha}$.

From definition (B.1) we conclude that the behaviour of the three parameter Mittag-Leffler function is the stretched exponential [74]

$$E^\delta_{\alpha,\beta}(-t^\alpha) \simeq \frac{1}{\Gamma(\beta)} - \delta\frac{t^\alpha}{\Gamma(\alpha+\beta)} \simeq \frac{1}{\Gamma(\beta)}\exp\left(-\delta\frac{\Gamma(\beta)}{\Gamma(\alpha+\beta)}t^\alpha\right). \tag{B.3}$$

Using the series expansion around $z=\infty$ [75] (for details see also [76])

$$E^\delta_{\alpha,\beta}(-z) = \frac{z^{-\delta}}{\Gamma(\delta)}\sum_{k=0}^{\infty}\frac{\Gamma(\delta+k)}{\Gamma(\beta-\alpha(\delta+n))}\frac{(-z)^{-n}}{n!}, \tag{B.4}$$

for $0<\alpha<2$ and $z\to\infty$, we find that the asymptotic behaviour of the three parameter Mittag-Leffler function is given by

$$E^\delta_{\alpha,\beta}(-t^\alpha) \simeq \frac{t^{-\alpha\delta}}{\Gamma(\beta-\alpha\delta)}, \quad t\to\infty. \tag{B.5}$$

The following formula for the derivative of the Mittag-Leffler function follows directly from definition (B.1) applying term-by-term differentiation,

$$\frac{d}{dt}\left(t^{\beta-1}E^\delta_{\alpha,\beta}\left(at^\alpha\right)\right) = t^{\beta-2}E^\delta_{\alpha,\beta-1}\left(at^\alpha\right). \tag{B.6}$$

From the generalised Langevin equation (22) and the exponentially truncated friction kernel (37) via the Laplace transform method, we find for the MSD

$$\left\langle x^2(t)\right\rangle = \frac{2k_BT}{m\Gamma_H}\mathcal{L}^{-1}\left[\frac{s^{-2}}{(s+\tau_\star^{-1})^{1-2H}}\right]. \tag{B.7}$$

Therefore, from the Laplace transform formula (B.2), where $\alpha \to 1$, $\delta \to 1-2H$, $\alpha\delta - \beta \to -2$, that is, $\beta \to 3-2H$, and $\nu \to \tau_\star^{-1}$, we obtain the result (39).

Appendix C. Derivations for section 4

Appendix C.1. Derivation of MSD for tfBm

Due to the white Gaussian noise in equation (64) the MSD of tempered fractional Brownian motion (64) can be written as

$$\begin{aligned}\left\langle B^2_{H,\lambda}(t)\right\rangle &= \sigma^2\Bigg[\int_0^t e^{-2\lambda(t-u)}(t-u)^{2H-1}du \\ &\quad + \int_{-\infty}^0 \left(e^{-\lambda(t-u)}(t-u)^{H-1/2} - e^{\lambda u}(-u)^{H-1/2}\right)^2 du\Bigg].\end{aligned} \tag{C.1}$$

After expanding the square of the second integral and using the appropriate changes of variable, it becomes

$$\left\langle B^2_{H,\lambda}(t)\right\rangle = \sigma^2\Big[\int_0^\infty e^{-2\lambda ts}s^{2H-1}ds - e^{-\lambda t}\int_0^\infty e^{-2\lambda ts}(1+s)^{H-1/2}s^{H-1/2}ds\Big]. \tag{C.2}$$

These integrals can be found, for instance, as equations (3.381 4) and (3.383 8) in [77]. This produces equation (65).

Appendix C.2. Derivation of autocorrelation function of tempered fractional Gaussian noise

In the classical paper by Mandelbrot and van Ness [36] a smooth fractional Brownian motion is defined in terms of the small and positive parameter δ, through

$$B_H(t;\delta) = \frac{1}{\delta}\int_t^{t+\delta} B_H(u)du. \tag{C.3}$$

Its derivative is known as the fractional Gaussian noise

$$B'_H(t;\delta) = \frac{1}{\delta}\left[B_H(t+\delta) - B_H(t)\right], \tag{C.4}$$

where we omit the explicit dependence on δ in the main text. The autocorrelation function of equation (C.4) is given in expression (72).

The same procedure can be applied to tempered fractional Brownian motion to define the corresponding continuous fractional noise

$$B'_{H,\lambda}(t;\delta) = \frac{1}{\delta}\left[B_{H,\lambda}(t+\delta) - B_{H,\lambda}(t)\right]. \tag{C.5}$$

With the identity

$$2(a-b)(c-d) = (a-d)^2 + (b-c)^2 - (a-c)^2 - (b-d)^2, \tag{C.6}$$

and the fact that tempered fractional Brownian motion has stationary increments, and $B_{H,\lambda}(0) = 0$, we obtain

$$\left\langle B'_{H,\lambda}(t;\delta)B'_{H,\lambda}(t+\tau;\delta)\right\rangle = \frac{1}{2\delta^2}\Big[\left\langle B^2_{H,\lambda}(\tau-\delta)\right\rangle + \left\langle B^2_{H,\lambda}(\tau+\delta)\right\rangle - 2\left\langle B^2_{H,\lambda}(\tau)\right\rangle\Big]. \quad \text{(C.7)}$$

By virtue of relation (65) the autocorrelation function of tempered fractional Gaussian noise becomes expression (70). The autocorrelation function of tempered fractional Gaussian noise (70) has a well defined limit when $\delta\lambda \to 0$,

$$\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle = \frac{\sigma^2\Gamma(H+1/2)\lambda^{2-2H}}{2^H\sqrt{\pi}}\Big[(\lambda|\tau|)^{H-1}K_{1-H}(\lambda|\tau|) - (\lambda|\tau|)^H K_{2-H}(\lambda|\tau|)\Big]. \quad \text{(C.8)}$$

Appendix C.3. Evaluating the integral over the autocorrelation function of fractional Gaussian noise

Taking the integral over expression (72) and denoting

$$W_H = \frac{\Gamma^2(H+\frac{1}{2})}{2\Gamma(2H+1)\sin(\pi H)} \quad \text{(C.9)}$$

one gets

$$\begin{aligned}
\mathbb{K} &= \int_0^\infty d\tau \lim_{\lambda\to 0}\left\langle B'_{H,\lambda}(t)B'_{H,\lambda}(t+\tau)\right\rangle \\
&= \frac{W_H}{\delta^2}\times\lim_{A\to\infty}\left[\int_0^A d\tau(\tau+\delta)^{2H} + \int_0^A d\tau|\tau-\delta|^{2H} - 2\int_0^A d\tau\tau^{2H}\right] \\
&= \frac{W_H}{\delta^2}\times\lim_{A\to\infty}\left[\int_\delta^{A+\delta} d\tau\tau^{2H} + \int_{-\delta}^{A-\delta} d\tau\tau^{2H} - 2\int_0^A d\tau\tau^{2H}\right] \\
&= \frac{W_H}{(2H+1)\delta^2}\times\lim_{A\to\infty}\left[(A+\delta)^{2H+1} - \delta^{2H+1} + (A-\delta)^{2H+1} - \delta^{2H+1} - 2A^{2H+1}\right] \\
&= \frac{W_H}{(2H+1)\delta^2}\times\lim_{A\to\infty}\left[A^{2H+1}\left(1+(2H+1)\frac{\delta}{A}+2H(2H+1)\frac{\delta^2}{2A^2}\right) - \delta^{2H+1}\right. \\
&\quad \left. + A^{2H+1}\left(1-(2H+1)\frac{\delta}{A}+2H(2H+1)\frac{\delta^2}{2A^2}\right) - \delta^{2H+1} - 2A^{2H+1}\right] \\
&= \frac{W_H}{(2H+1)\delta^2}\times\lim_{A\to\infty}\left[2H(2H+1)\delta^2A^{2H-1}\right] \\
&= \frac{\Gamma^2(H+\frac{1}{2})}{2\Gamma(2H)\sin(\pi H)}\times\lim_{A\to\infty}\left[A^{2H-1}\right] \\
&= \begin{cases} \infty, & H > \frac{1}{2} \\ 0, & H < \frac{1}{2} \end{cases}. \quad \text{(C.10)}
\end{aligned}$$

Appendix C.4. Tempered fractional Gaussian noise: MSD for short observation times

For the regime of short observation times, $\delta \ll t \ll \lambda^{-1}$, we apply the linear transformation 15.3.6 from [42] for hypergeometric functions. In the resulted definition,

the argument of the hypergeometric function is small,

$$\tilde{\gamma}_H(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\left\{\frac{\pi 2^{H-1}s}{\sin(\pi H)\Gamma(1-H)\lambda^H} - s^2\frac{s^{-1}\lambda^{-H}}{2^{H+1}}\sqrt{\pi}\Gamma\left[\begin{array}{c}1,2H+1\\H+\frac{3}{2}\end{array}\right]\right.$$
$$\times\left[\frac{\Gamma(H+\frac{3}{2})\Gamma(H)}{\Gamma(H+1)\Gamma(H+\frac{1}{2})}\,{}_2F_1\left(\frac{1}{2},1;1-H;\frac{\lambda^2}{s^2}\right)\right.$$
$$\left.\left.+\left(\frac{\lambda^2}{s^2}\right)^H\frac{\Gamma(H+\frac{3}{2})\Gamma(-H)}{\Gamma(\frac{1}{2})\Gamma(1)}\,{}_2F_1\left(H+1,H+\frac{1}{2};H+1;\frac{\lambda^2}{s^2}\right)\right]\right\}. \tag{C.11}$$

For small arguments we use the general definition of hypergeometric functions, 15.1.1 in [42], up to the second order. Then

$$\tilde{\gamma}_H(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\left\{\frac{\pi 2^{H-1}s}{\sin(\pi H)\Gamma(1-H)\lambda^H} - \frac{\sqrt{\pi}s}{2^{H+1}\lambda^H}\Gamma\left[\begin{array}{c}1,2H+1\\H+\frac{3}{2}\end{array}\right]\right.$$
$$\times\left[\frac{\Gamma\left(H+\frac{3}{2}\right)\Gamma(H)}{\Gamma(H+1)\Gamma\left(H+\frac{1}{2}\right)}\frac{\Gamma(1-H)}{\Gamma\left(\frac{1}{2}\right)}\left(\frac{\Gamma(\frac{1}{2})}{\Gamma(1-H)}+\frac{\Gamma(1+\frac{1}{2})\Gamma(2)}{\Gamma(2-H)}\frac{\lambda^2}{s^2}\right)\right.$$
$$+\left(\frac{\lambda^2}{s^2}\right)^H\frac{\Gamma(H+\frac{3}{2})\Gamma(-H)}{\Gamma(\frac{1}{2})}\frac{\Gamma(H+1)}{\Gamma\left(H+\frac{1}{2}\right)\Gamma(H+1)}\left(\Gamma\left(H+\frac{1}{2}\right)\right.$$
$$\left.\left.\left.+\Gamma\left(H+\frac{3}{2}\right)\frac{\lambda^2}{s^2}\right)\right]\right\}. \tag{C.12}$$

Now, we simplify the Gamma functions using the duplication formula 6.1.18 in [42],

$$\tilde{\gamma}_H(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\left\{\frac{\pi 2^{H-1}s}{\sin(\pi H)\Gamma(1-H)\lambda^H} - \frac{2^{H-1}\Gamma(H)s}{\lambda^H} - \frac{2^{H-1}\Gamma(H)s}{\lambda^H 2(1-H)}\frac{\lambda^2}{s^2}\right.$$
$$-\frac{2^{H-1}s}{\lambda^H}\frac{H}{(H+\frac{1}{2})}\frac{\Gamma(H)\Gamma(-H)\Gamma(H+\frac{3}{2})}{\Gamma(\frac{1}{2})}\left(\frac{\lambda^2}{s^2}\right)^H$$
$$\left.-\frac{2^{H-1}s}{\lambda^H}\frac{H}{(H+\frac{1}{2})}\frac{\Gamma(H)\Gamma(-H)\Gamma(H+\frac{3}{2})}{\Gamma(\frac{1}{2})}(H+\frac{1}{2})\left(\frac{\lambda^2}{s^2}\right)^{H+1}\right\} \tag{C.13}$$

Using Euler's reflection formula,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)} \tag{C.14}$$

the first two terms cancel each other and it can be seen that the dominant term in the autocorrelation function scales as s^{1-2H},

$$\tilde{\gamma}_H(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\left\{\frac{2^{H-1}}{\lambda^{H-1}(H+\frac{1}{2})}\frac{\pi}{\sin(\pi H)}\frac{\Gamma(H+\frac{3}{2})}{\sqrt{\pi}}\left(\frac{\lambda}{s}\right)^{2H-1}\right.$$
$$\left.-\frac{2^{H-2}\Gamma(H)}{\lambda^{H-1}(1-H)}\left(\frac{\lambda}{s}\right)+\frac{2^{H-1}}{\lambda^{H-1}}\frac{\pi\Gamma(H+\frac{3}{2})}{\sin(\pi H)\sqrt{\pi}}\left(\frac{\lambda}{s}\right)^{2H+1}\right\}. \tag{C.15}$$

Appendix C.5. Tempered fractional Gaussian noise: MSD for long observation time

For the regime of long observation time or $\frac{\lambda}{s} \gg 1$, we go back to equation (82) and use relation (15.3.8) from [42] for hypergeometric functions with small arguments. Then,

by applying the expansion of hypergeometric functions up to the second order for small argument, $s/\lambda \ll 1$,

$$
\begin{aligned}
&{}_2F_1\left(\frac{1}{2},1;H+\frac{3}{2};1-\frac{\lambda^2}{s^2}\right) = \left(\frac{s}{\lambda}\right)\frac{\Gamma(H+\frac{3}{2})\Gamma(\frac{1}{2})}{\Gamma(1)\Gamma(H+1)}\,{}_2F_1\left(\frac{1}{2},H+\frac{1}{2};\frac{1}{2};\frac{s^2}{\lambda^2}\right)\\
&+\left(\frac{s^2}{\lambda^2}\right)\frac{\Gamma(H+\frac{3}{2})\Gamma(-\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(H+\frac{1}{2})}\,{}_2F_1\left(1,H+1;\frac{3}{2};\frac{s^2}{\lambda^2}\right)\\
&=\left(\frac{s}{\lambda}\right)\frac{\Gamma(H+\frac{3}{2})\sqrt{\pi}}{\Gamma(H+1)}\left[\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(H+\frac{1}{2})}\sum_{k=0}^{\infty}\frac{\Gamma(k+\frac{1}{2})\Gamma(k+H+\frac{1}{2})}{\Gamma(k+\frac{1}{2})}\frac{1}{k!}\left(\frac{s^2}{\lambda^2}\right)^k\right]\\
&+\left(\frac{s^2}{\lambda^2}\right)(H+\frac{1}{2})(-2)\left[\frac{\Gamma(\frac{3}{2})}{\Gamma(1)\Gamma(H+1)}\sum_{k=0}^{\infty}\frac{\Gamma(k+1)\Gamma(k+H+1)}{\Gamma(k+\frac{3}{2})}\frac{1}{k!}\left(\frac{s^2}{\lambda^2}\right)^k\right]\\
&=\frac{s}{\lambda}\frac{\left(H+\frac{1}{2}\right)\sqrt{\pi}}{\Gamma(H+1)}\left[\Gamma\left(H+\frac{1}{2}\right)+\Gamma\left(H+\frac{3}{2}\right)\frac{s^2}{\lambda^2}\right]\\
&-\frac{s^2}{\lambda^2}\frac{\left(H+\frac{1}{2}\right)\sqrt{\pi}}{\Gamma(H+1)}\left[\frac{\Gamma(H+1)}{\Gamma\left(\frac{3}{2}\right)}+\frac{\Gamma(2)\Gamma(H+2)}{\Gamma\left(\frac{3}{2}+1\right)}\frac{s^2}{\lambda^2}\right].
\end{aligned}
\tag{C.16}
$$

As a result, the integral in expression (82) is approximated as

$$
\begin{aligned}
\int_0^\infty dt e^{-st}t^H K_H(\lambda t) \sim s^{-1}2^{H-1}\lambda^{-H}\Bigg\{&\sqrt{\pi}\Gamma(H+\frac{1}{2})\frac{s}{\lambda}\\
&+\sqrt{\pi}\Gamma(H+\frac{3}{2})\frac{s^3}{\lambda^3}-2\Gamma(H+1)\frac{s^2}{\lambda^2}-\frac{4}{3}(H+1)\Gamma(H+1)\frac{s^4}{\lambda^4}\Bigg\}
\end{aligned}
\tag{C.17}
$$

Applying these approximations, the resulting expression for the autocorrelation function in the Laplace domain is

$$
\begin{aligned}
\tilde{\gamma}(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\Bigg\{&\frac{\pi 2^{H-1}s}{\sin(\pi H)\Gamma(1-H)\lambda^H}-s^2s^{-1}2^{H-1}\lambda^{-H}\\
&\times\left[\sqrt{\pi}\Gamma(H+\frac{1}{2})\frac{s}{\lambda}+\sqrt{\pi}\Gamma(H+\frac{3}{2})\frac{s^3}{\lambda^3}-2\Gamma(H+1)\frac{s^2}{\lambda^2}\right.\\
&\left.-\frac{4}{3}(H+1)\Gamma(H+1)\frac{s^4}{\lambda^4}\right]\Bigg\}.
\end{aligned}
\tag{C.18}
$$

It can be seen that the dominant term is a linear function of s,

$$
\tilde{\gamma}(s) = \frac{2\Gamma(H+\frac{1}{2})}{\sqrt{\pi}(2\lambda)^H}\frac{\pi 2^{H-1}s}{\sin(\pi H)\Gamma(1-H)\lambda^H}
\tag{C.19}
$$

Appendix D. Derivation of the MSD of the fractional Ornstein-Uhlenbeck process

The solution of equation (89) for a general noise $\xi(u)$ is

$$
x(t) = e^{-\lambda t}\int_0^t e^{\lambda u}\xi(u)\,du,
\tag{D.1}
$$

so

$$
\left\langle x^2(t)\right\rangle = e^{-2\lambda t}\int_0^t\int_0^t e^{\lambda(u_1+u_2)}\langle\xi(u_1)\xi(u_2)\rangle du_1 du_2.
\tag{D.2}
$$

In general, for a noise such that $\langle \xi(u_1)\xi(u_2)\rangle = g(|u_1 - u_2|)$, equation (D.2) becomes

$$\left\langle x^2(t)\right\rangle = \frac{1}{\lambda}\left[\int_0^t e^{-\lambda\tau} g(\tau)d\tau - e^{-2\lambda t}\int_0^t e^{\lambda\tau} g(\tau)d\tau\right]. \tag{D.3}$$

In our case, $\xi(u) = B'_H(u)$. For $H \neq 1/2$, $g(u) = \sigma^2 H(2H-1)V_H u^{2H-2}$ and the MSD can be expressed in terms of the Kummer function $M(a;b;z)$,

$$\begin{aligned}\left\langle x^2(t)\right\rangle \;=\; & \frac{\sigma^2 H V_H t^{2H-1}}{\lambda}\Big[M(2H-1;2H;-\lambda t) \\ & - e^{-2\lambda t}M(2H-1;2H;\lambda t)\Big].\end{aligned} \tag{D.4}$$

If $H = 1/2$, using $g(u) = \sigma^2\delta(u)$ in equation (D.2), we arrive at

$$\langle x^2(t)\rangle = \frac{\sigma^2}{2\lambda}\left(1 - e^{2\lambda t}\right). \tag{D.5}$$

This result coincides with equation (D.4) for $H = 1/2$, such that equation (D.4) is valid for all $H \in (0,1)$. Using the properties of the Kummer function (which in our case reduces to the incomplete gamma function), relation (D.4) is shown to be equivalent to equation (90).

Acknowledgments

This research is supported by the Basque Government through the BERC 2014-2017 and BERC 2018-2021 programmes and by Spanish Ministry of Economy and Competitiveness MINECO, BCAM Severo Ochoa excellence accreditation SEV-2013-0323, and project MTM2016-76016-R "MIP". TS, AC and RM acknowledge funding from the Deutsche Forschungsgemeinschaft, project ME 1535/6-1. RM acknowledges support from Deutsche Forschungsgemeinschaft, project ME 1535/7-1, as well as from the Foundation for Polish Science (Fundacja na rzecz Nauki Polski) in the framework of a an Alexander von Humboldt Polish Honorary Research Fellowship.

References

[1] R. Brown. A brief account of microscopical observations made on the particles contained in the pollen of plants. Phil. Mag. **4**, 161 (1828).

[2] J. Perrin. L'agitation moléculaire et le mouvement brownien. Compt. Rend. (Paris) **146**, 967 (1908).

[3] I. Nordlund. Eine neue Bestimmung der Avogadroschen Konstante aus der Brownschen Bewegung kleiner, in Wasser suspendierten Quecksilberkügelchen. Z. Phys. Chem. **87**, 40 (1914).

[4] E. Kappler. Versucher zur Messung der Avogadro-Loschmidtschen Zahl aus der Brownschen Bewegung einer Drehwaage. Ann. Phys. (Leipzig) **11**, 233 (1932).

[5] J.-P. Bouchaud and A. Georges. Anomalous diffusion in disordered media: statistical mechanisms, models, and physical applications. Phys. Rep. **195**, 127 (1990).

[6] R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in fractional dynamics descriptions of anomalous dynamical processes. Phys. Rep. **339**, 1 (2000).

[7] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. Phys. Chem. Chem. Phys. **16**, 24128 (2014).

[8] F. Höfling and T. Franosch. Anomalous transport in the crowded world of biological cells. Rep. Prog. Phys. **76**, 046602 (2013).

[9] K. Nørregaard, R. Metzler, C. Ritter, K. Berg-Sørensen, and L. Oddershede. Manipulation and motion of organelles and single molecules in living cells. Chem. Rev. **117**, 4342 (2017).

[10] A. Caspi, R. Granek, and M. Elbaum. Enhanced diffusion in active intracellular transport. Phys. Rev. Lett. **85**, 5655 (2000).

[11] G. Seisenberger, M. U. Ried, T. Endreß, H. Büning, M. Hallek, and C. Bräuchle. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. Science **294**, 1929 (2001).

[12] M. Weiss, M. Elsner, F. Kartberg, and T. Nilsson. Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. Biophys. J. **87**, 3518 (2004).

[13] I. M. Tolić-Nørrelykke, E.-L. Munteanu, G. Thon, L. Oddershede, and K. Berg-Sørensen. Anomalous diffusion in living yeast cells. Phys. Rev. Lett. **93**, 078102 (2004).

[14] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai and Y. Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. Phys. Rev. Lett. **103**, 018102 (2009).
K. Burnecki, E. Kepten, J. Janczura, I. Bronshtein, Y. Garini, and A. Weron. Universal algorithm for identification of fractional Brownian motion. A case of telomere subdiffusion. Biophys. J. **103**, 1839 (2012).

[15] J.-H. Jeon, V. Tejedor, S. Burov, E. Barkai, C. Selhuber-Unke, K. Berg-Soerensen, L. Oddershede, and R. Metzler, In vivo anomalous diffusion and weak ergodicity breaking of lipid granules. Phys. Rev. Lett. **106**, 048103 (2011).

[16] S. M. A. Tabei, S. Burov, H. Y. Kim, A. Kuznetsov, T. Huynh, J. Jureller, L. H. Philipson, A. R. Dinner, and N. F. Scherer. Intracellular transport of insulin granules is a subordinated random walk. Proc. Natl. Acad. Sci. USA **110**, 4911 (2013).

[17] T. J. Lampo, S. Stylianido, M. P. Backlund, P. A. Wiggins, and A. J. Spakowitz. Cytoplasmic RNA-protein particles exhibit non-Gaussian subdiffusive behaviour. Biophys. J. **112**, 532 (2017).
R. Metzler. Gaussianity fair: the riddle of anomalous yet non-Gaussian diffusion. Biophys. J. **112**, 413 (2017).

[18] D. S. Banks and C. Fradin. Anomalous diffusion of proteins due to molecular crowding. Biophys. J. **89**, 2960 (2005).

[19] J. Szymanski and M. Weiss. Elucidating the origin of anomalous diffusion in crowded fluids. Phys. Rev. Lett. **103**, 038102 (2009).

[20] J.-H. Jeon, N. Leijnse, L. B. Oddershede, and R. Metzler. Anomalous diffusion and power-law relaxation in wormlike micellar solution. New J. Phys. **15**, 045011 (2013).

[21] X. Hu, L. Hong, M. D. Smith, T. Neusius, X. Cheng, and J. C. Smith. The dynamics of single protein molecules is non-equilibrium and self-similar over thirteen decades in time. Nature Phys. **12**, 171 (2016).

[22] J.-H. Jeon, H. M.-S. Monne, M. Javanainen, and R. Metzler. Lateral motion of phospholipids and cholesterols in a lipid bilayer: anomalous diffusion and its origins. Phys. Rev. Lett. **109**, 188103 (2012).

[23] M. Javanainen, H. Hammaren, L. Monticelli, J.-H. Jeon, M. S. Miettinen, H. Martinez-Seara, R. Metzler, and I. Vattulainen. Anomalous and normal diffusion of proteins and lipids in crowded lipid membranes. Faraday Disc. **161**, 397 (2013).

[24] J.-H. Jeon, M. Javanainen, H. Martinez-Seara, R. Metzler, and I. Vattulainen. Protein crowding in lipid bilayers gives rise to non-Gaussian anomalous lateral diffusion of phospholipids and proteins. Phys. Rev. X **6**, 021006 (2016).

[25] G. R. Kneller, K. Baczynski, and M. Pasienkewicz-Gierula. Consistent picture of lateral subdiffusion in lipid bilayers: molecular dynamics simulation and exact results. J. Chem. Phys. **135**, 141105 (2011).

[26] S. Stachura and G. R. Kneller. A scaling approach to anomalous diffusion. J. Chem. Phys. **40**, 245

(2014).
[27] K. Chen, B. Wang, and S. Granick. Memoriless self-reinforcing directionality in endosomal active transport within living cells. Nature Mat. **14**, 589 (2015).
[28] J. F. Reverey, J.-H. Jeon, M. Leippe, R. Metzler, and C. Selhuber-Unkel. Superdiffusion dominates intracellular particle motion in the supercrowded space of pathogenic Acanthamoeba castellanii. Sci. Rep. **5**, 11690 (2015).
[29] M. S. Song, H. C. Moon, J.-H. Jeon, and H. Y. Park. Neuronal messenger ribonucleoprotein transport follows an aging Lévy walk. Nat. Comm. **9**, 344 (2018).
[30] I. Vattulainen and T. Róg. Lipid membranes: Theory and simulations bridged to experiments. Biochimica et Biophysica Acta (BBA) - Biomembranes **1858**, 2251 (2016).
[31] R. Metzler, J.-H. Jeon, and A. G. Cherstvy. Non-Brownian diffusion in lipid membranes: Experiments and simulations. Biochimica et Biophysica Acta (BBA) - Biomembranes **1858**, 2451 (2016).
[32] Y. Meroz and I. M. Sokolov. A toolbox for determining subdiffusive mechanisms. Phys. Rep. **573**, 1 (2015).
[33] I. M. Sokolov. Models of anomalous diffusion in crowded environments. Soft Matter **8**, 9043 (2012).
[34] H. Risken, The Fokker-Planck equation (Springer, Heidelberg, 1989).
[35] N. G. van Kampen, Stochastic processes in physics and chemistry (North Holland, Amsterdam, 1981).
[36] B. B. Mandelbrot and J. W. van Ness. Fractional brownian motions, fractional noises and applications. SIAM Rev. **10**, 422 (1968).
[37] I. M. Gel'fand and G. E. Shilov, Generalized Functions, Volume 1 (Academic Press, New York and London, 1964).
[38] R. Zwanzig, Nonequilibrium Statistical Mechanics (Oxford University Press, New York, 2001).
[39] W. T. Coffey and Y. P. Kalmykov, The Langevin equation (World Scientific, Singapore, 2012).
[40] Yu. L. Klimontovich, Turbulent motion and the structure of chaos (Kluwer, Dordrecht, 1991).
[41] Norton, M. P. & Karczub, D. G. Fundamentals of noise and vibration analysis for engineers (Cambridge University Press, Cambridge UK, 2003).
[42] M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables (Dover Books on Mathematics)", (Dover Publications, 1965).
[43] A. Erdelyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi, *Higher Transcedential Functions*, Vol. 3 (McGraw-Hill, New York, 1955).
[44] A. Mallet, Numerical inversion of Laplace transform, (Wolfram Library Archive, Item 0210–968, 2000).
[45] R. Kubo. The fluctuation-dissipation theorem. Rep. Prog. Phys. **29**, 255 (1966).
[46] W. Deng and E. Barkai. Ergodic properties of fractional Brownian-Langevin motion. Phys. Rev. E **79**, 011112 (2009).
[47] E. Lutz. Fractional Langevin equation. Phys. Rev. E **64**, 051106 (2001).
[48] I. Goychuk. Viscoelastic subdiffusion: generalized Langevin equation approach. Adv. Chem. Phys. **150**, 187 (2012).
[49] N. Pottier. Aging properties of an anomalously diffusing particule. Physica A **317**, 371 (2003).
[50] A. Liemert, T. Sandev, and H. Kantz. Generalized Langevin equation with tempered memory kernel. Physica A **466**, 356 (2017).
[51] F. Durbin. Numerical inversion of Laplace transforms. Comp. J. **17**, 371 (1974).
[52] C. Beck. Dynamical Foundations of Nonextensive Statistical Mechanics. Phys. Rev. Lett. **87** 180601 (2001).
E. van der Straeten and C. Beck. Superstatistical distributions from a maximum entropy principle. Phys. Rev. E **78**, 051101 (2008).
E. van der Straeten and C. Beck. Dynamical modelling of superstatistical complex systems. Physica A **390**, 951 (2011).
J. Ślezak, R. Metzler, and M. Magdziarz. Superstatistical generalised Langevin equation: non-

Gaussian viscoelastic anomalous diffusion New J. Phys. **20**, 023026 (2018).

[53] M. V. Chubynsky and G. W. Slater. Diffusing diffusivity: a model for anomalous, yet Brownian, diffusion. Phys. Rev. Lett. **113**, 098300 (2014).
R. Jain and K. L. Sebastian. Diffusion in a Crowded, Rearranging Environment. J. Phys. Chem. B **120** 3988 (2016).
A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov. Brownian yet non-Gaussian diffusion: from superstatistics to subordination of diffusing diffusivities Phys. Rev. X **7**, 021002 (2017).

[54] V. Sposini, A. V. Chechkin, F. Seno, G. Pagnini, and R. Metzler. Random diffusivity from stochastic equations: comparison of two models for Brownian yet non-Gaussian diffusion. New J. Phys. **20**, 043044 (2018).

[55] M. Javanainen H. Martinez-Seara R. Metzler, and I. Vattulainen. Diffusion of Integral Membrane Proteins in Protein-Rich Membranes. J. Phys. Chem. Lett. **8**, 4308 (2017).

[56] M. M. Meerschaert and F. Sabzikar. Tempered fractional Brownian motion. Statistics & Probability Lett., **83**, 2269 (2013).

[57] Y. Chen, X. Wang and W. Deng. Localization and ballistic diffusion for the tempered fractional Brownian-Langevin motion. J. Stat. Phys. **169**, 18 (2017).

[58] Yu. A. Brychkov, O. I. Marichev and A. P. Prudnikov, "Integrals and Series. Volume 2. Special Features / Integraly i ryady. Tom 2. Spetsialnye funktsii", (FIZMATLIT, 2003).

[59] O. Yu. Sliusarenko, V. Yu. Gonchar, A. V. Chechkin, I. M. Sokolov, and R. Metzler. Kramers escape driven by fractional Brownian motion. Phys. Rev. E **81**, 041119 (2010).

[60] J.-H. Jeon and R. Metzler. Inequivalence of time and ensemble averages in ergodic systems: exponential versus power-law relaxation in confinement. Phys. Rev. E **85**, 021147 (2012).

[61] M. Schwarzl, A. Godec, and R. Metzler. Quantifying non-ergodicity of anomalous diffusion with higher order moments. Sci. Rep. **7**, 3878 (2017).

[62] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, Integrals and Series, Vol. 1 (Taylor & Francis, London, 2002).

[63] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series*, Vol. 3 (Gordon and Breach Science Publishers, New York, 1992).

[64] C. Carretero-Campos, P. Bernaola-Galván, P. Ch. Ivanov, and P. Carpena. Phase transitions in the first-passage time of scale-invariant correlated processes. Phys. Rev. E **85**, 011139 (2012).

[65] P. Carpena, A. V. Coronado, C. Carretero-Campos, P. Bernaola-Galván, and P. Ch. Ivanov. First-Passage Time Properties of Correlated Time Series with Scale-Invariant Behavior and with Crossovers in the Scaling. In Time Series Analysis and Forecasting, edited by I. Rojas and H. Pomares (Springer, Berlin, 2016)

[66] H. Scher and E. W. Montroll. Anomalous transit-time dispersion in amorphous solids. Phys. Rev. B **12**, 2455 (1975).

[67] G. Pagnini. Short note on the emergence of fractional kinetics. Physica A **409**, 29 (2014).

[68] J.-H. Jeon, A. V. Chechkin, and R. Metzler. Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion. Phys. Chem. Chem. Phys. **16**, 15811 (2014).

[69] A. G. Cherstvy, A. V. Chechkin, and R. Metzler. Non-ergodicity, fluctuations, and criticality in heterogeneous diffusion processes. New J. Phys. **15**, 083039 (2013).

[70] A. Mura and G. Pagnini. Characterizations and simulations of a class of stochastic processes to model anomalous diffusion. J. Phys. A **41**, 285003 (2008).
D. Molina-García, T. Minh Pham, P. Paradisi, C. Manzo, and G. Pagnini. Fractional kinetics emerging from ergodicity breaking in random media. Phys. Rev. E **94**, 052147 (2016).

[71] S. Vitali, V. Sposini, O. Sliusarenko, P. Paradisi, G. Castellani, and G. Pagnini. Langevin equation in complex media and anomalous diffusion. E-print arXiv:1806.11508.

[72] R. Schumer, A. Taloni, and D. J. Furbish. Theory connecting nonlocal sediment transport, earth surface roughness, and the Sadler effect. Geophys. Res. Lett. **44**, 2281 (2017).

[73] T. R. Prabhakar. A Singular Equation with a Generalized Mittag-Leffler Function in the Kernel.

Yokohama Math. J. **19**, 7 (1971).

[74] T. Sandev, A.V. Chechkin, N. Korabel, H. Kantz, I. M. Sokolov, and R. Metzler. Distributed-order diffusion equations and multifractality: Models and solutions. Phys. Rev. E **92**, 042117 (2015).

[75] T. Sandev, A. Chechkin, H. Kantz, and R. Metzler. Diffusion and Fokker-Planck-Smoluchowski Equations with Generalized Memory Kernel. Fract. Calc. Appl. Anal. **18**, 1006 (2015).

[76] R. Garra and R. Garrappa. The Prabhakar or three parameter Mittag–Leffler function: theory and application. Comm. Nonlin. Sci. Numer. Simul. **56**, 314 (2018).

[77] I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series, and Products, Seventh edition (Academic Press, 2007).

Appendix C

Article 3

The reference of this contribution is:

Pigolotti, S.; Cencini, M.; Molina, D. & Muñoz, M. A.
Stochastic spatial models in ecology: a statistical physics approach
Journal of Statistical Physics, Springer, **2018**, 172, 44-73
https://doi.org/10.1007/s10955-017-1926-4

SJR 2017: 0.930
Relative position: #14/49 (Percentile 72%)
Mathematics: Mathematical Physics

Stochastic spatial models in ecology: a statistical physics approach

Simone Pigolotti[1] Massimo Cencini[2], Daniel Molina[3] Miguel A. Muñoz[4]
[1] *Biological Complexity Unit, Okinawa Institute of Science and Technology and Graduate University, Onna, Okinawa 904-0495.* [2] *Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, via dei Taurini 19, 00185 Rome, Italy.* [3] *BCAM - Basque Center for Applied Mathematics, Alameda de Mazarredo 14, E-48009 Bilbao, Basque Country, Spain.* [4] *Departamento de Electromagnetismo y Física de la Materia, and Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, 18071 Granada, Spain.*

Ecosystems display a complex spatial organization. Ecologists have long tried to characterize them by looking at how different measures of biodiversity change across spatial scales. Ecological neutral theory has provided simple predictions accounting for general empirical patterns in communities of competing species. However, while neutral theory in well-mixed ecosystems is mathematically well understood, spatial models still present several open problems, limiting the quantitative understanding of spatial biodiversity. In this review, we discuss the state of the art in spatial neutral theory. We emphasize the connection between spatial ecological models and the physics of non-equilibrium phase transitions and how concepts developed in statistical physics translate in population dynamics, and vice versa. We focus on non-trivial scaling laws arising at the critical dimension $D = 2$ of spatial neutral models, and their relevance for biological populations inhabiting two-dimensional environments. We conclude by discussing models incorporating non-neutral effects in the form of spatial and temporal disorder, and analyze how their predictions deviate from those of purely neutral theories.

I. INTRODUCTION

Community ecology aims at shedding light on how competing species assemble and coexist in their habitats [1]. This has proven to be a formidable challenge. A main reason is that ecological dynamics span a wide range of spatial and temporal scales, from those typical of individuals to those characterizing large populations or communities. Ecologists have empirically characterized biodiversity at the different spatial scales; for example, counting the average number of species hosted in a given area – *species area relationship* (SAR) [2, 3]–, or the distribution of their abundances – *species abundance distribution* (SAD) [4, 5]. Often, the ecological forces determining these patterns act at a given spatio-temporal scale but can affect others as well. The inverse problem, i.e. linking observed patterns with the causes originating them at different scales, is arguably the central problem in ecology [6].

This kind of problem sounds familiar to experts in statistical physics, where large-scale emergent behavior results from interactions among simple local units. Tools of statistical physics are indeed very useful to make progress on the aforementioned crucial issues in ecology. In particular, a natural approach to such complex problems is to radically simplify them. To this aim, we consider ecosystems made up of competing non-motile species, such as trees, or having a motility range much smaller than the typical linear size of the population, such as communities of microorganisms. Further possible simplifications are that all emergent phenomena originate at the single-individual scale and, more drastically, that differences among individuals, possibly belonging to different species, can be neglected. These assumptions constitute the basis of the *ecological neutral theory* proposed by Hubbell [7].

Ecological neutral theory [7] was built upon theoretical ideas of Kimura's neutral theory of population genetics [8]. Both theories underscore the role of stochastic demographic fluctuations in determining the fate of populations and completely neglect deterministic effects stemming from fitness differences. The assumption of ecological neutrality has elicited heated controversies, as it hinted that classical ecological concepts, such as niches, might play a marginal role in structuring communities of competing species. Despite these contentions, neutral theory had a considerable impact on ecological thinking, owing to its ability to quantitatively predict non-trivial patterns of biodiversity with simple models characterized by very few adjustable parameters [9–11].

Spatially implicit neutral models describe well-mixed communities of individuals subject to immigration from a larger reservoir of species where diversity is maintained via speciation. They can be solved analytically [12–16], yielding analytical expressions for the SAD. Beside the mathematical appeal, these exact solutions have been extremely helpful for fitting empirical data and therefore testing neutral theory or, at least, promote it as a null-model [17]. For more exhaustive surveys of ecological neutral theory, we refer the reader to Hubbell's book [7] and the reviews [9–11].

The focus of this review is on spatially-explicit neutral and near-neutral population models. Explicitly describing space is crucial to address the fundamental ecological questions sketched at the beginning of the introduction. However, spatially-explicit models – that are often variants of familiar models in non-equilibrium statistical

physics [18] – are still poorly understood, especially if compared with their well-mixed counterparts [19]. One of the most studied neutral model is the *voter model with speciation*, or multi-species voter model [20–22], which generalizes the more common two-species voter model [23]. The stepping-stone model [24–27] and the contact process [23, 28, 29] are other examples of spatial models that have been studied in both the physics and population biology literature. We shall discuss how these analogies can be used to advance our understanding of spatial ecology and the main open problems. This review heavily relies on extensive numerical computations of lattice models based on previous works by the authors. This might have biased the choice of some topics and we apologize if some relevant works are not properly discussed.

The review is organized as follows. In Sect. II we introduce the multispecies voter model on a lattice and its dual representation in terms of coalescing random walkers. We then discuss its predictions of macroecological patterns: the SAR, and the SAD. For the latter, we compare two recent analytical approaches [30–32] with novel computational results. We mainly discuss the two-dimensional case due to its ecological relevance, but also briefly present the one-dimensional case for comparison. We conclude the section by presenting new results on an important dynamical property: the distribution of species persistence-times. In Sect. III we discuss other neutral models, where, at variance with the voter model, lattice sites are not necessarily occupied by exactly one individual at all times. In particular, we consider the stepping stone model, where each lattice site hosts a local community of individuals. This generalization is relevant for modeling microorganisms and their macroecological patterns. We then consider a multispecies variant of the contact process, where lattice sites can be either empty of occupied by single individual. In Section IV we introduce non-neutral effects on a simplified two-species competition model, where adjusting a single parameter one can tune the departure from neutrality, here modeled as a specific habitat preference. Physically, this habitat preference can be thought as a form of quenched disorder. We discuss how this disorder generically favors species coexistence using the language of statistical mechanics, and also discuss other forms of disorder such as temporal heterogeneity. Finally, Sect. V is devoted to perspectives and conclusions.

II. VOTER MODEL WITH SPECIATION

A. Description of the model

A paradigmatic example of spatial neutral model is the *voter model with speciation*, [20], which is is a multi-species generalization of the voter model [23]. The latter is a widely studied model that has been applied in diverse contexts, from population genetics to spatial conflicts [33], spreading of epidemic diseases [34], opinion dynamics [35] and linguistics [36].

The voter model with speciation is defined on a lattice, where each site hosts one individual belonging to some species. At each discrete time step, a lattice site is chosen at random and the residing individual is removed (death event). Then, as illustrated Fig. 1, the dead individual is replaced:

- With probability ν, by an individual of a new species not present in the system (speciation event). Notice that, because of speciation, the total number of species is not fixed. In population genetics, this type of event is interpreted as a mutation within the same species [25, 37].
- With complementary probability $(1-\nu)$, by a new individual of an existing species (reproduction event). In this case, the newborn belongs to the same species of a parent individual chosen at random in the neighborhood of the vacant site. In the simplest case, the nearest-neighbors (NN) are chosen with uniform probability. More generally, the parent individual is selected according to a probability distribution $P(\vec{r})$ (the *dispersal kernel*) over the neighbors within a distance $\vec{r}$.

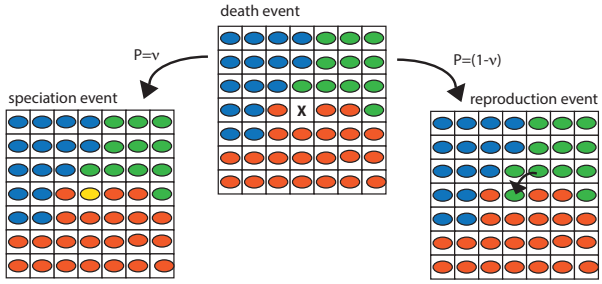


FIG. 1: Examples of transitions in the $2D$ voter model with speciation.

Most of this section will be devoted to the ecologically relevant case where the system is a two-dimensional ($2D$) square lattice, although we will briefly present some results in $1D$ for comparison.

B. Duality

The voter model with speciation is *dual* to a system of coalescing random walkers with an annihilation rate [20, 38, 39]. In this context, "duality" means that each trajectory of one system can be mapped in one of the other system having equal probability [38]. The dual process is constructed as follows. We start by placing on each lattice site a random walker. The dynamic of the dual process proceeds backward in time. At each discrete (backward) time step, with probability $1-\nu$, a randomly chosen walker is moved to a new site, where

the dispersal kernel $P(\vec{r})$ here plays the role of the distribution of possible displacements. If the site is occupied, the two walkers coalesce, i.e. one of the two is removed keeping trace of the coalescing partner. With complementary probability ν a randomly chosen random walker is annihilated, i.e. removed from the system. This event corresponds to a speciation event in the forward dynamics. The whole tree of coalescing random walkers, before annihilation, represents the entire genealogical tree of a species up to the speciation event that originated it.

The standard forward in time evolution of the voter-model with speciation and its dual dynamics are sketched, for the one-dimensional case, in Fig. 2a and 2b, respectively.

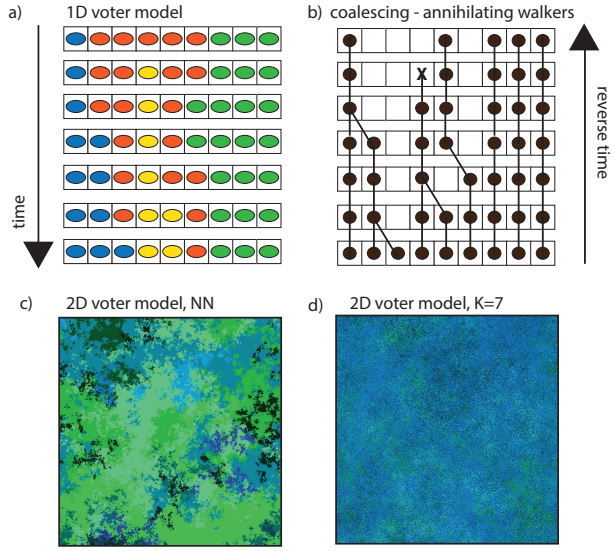


FIG. 2: a) Example of space-time dynamics of the $1D$ voter model with speciation. b) Corresponding dual dynamics: coalescing and annihilating random walkers. c) Snapshot of a configuration of the $2D$ voter model simulated with the dual dynamics, with $\nu = 5\ 10^{-7}$ and nearest-neighbor (NN) dispersal. d) Same as c) but with a longer dispersal range (uniformly distributed in a square of side K) with $K = 7$. Each color labels a different species.

Duality is a very useful property to understand the physics of the voter model. For example, it immediately stems from duality that the $\nu \to 0$ limit is fundamentally different in $D \leq 2$ and $D > 2$. As a matter of fact, in $D \leq 2$ the random walk is recurrent, meaning that the probability of two randomly chosen individuals to belong to the same species approaches one as $\nu \to 0$. In other words, in the absence of speciation, one has monodominance of one species in the long term. The same property does not hold in $D > 2$, where random walkers are not recurrent and, in an infinite system, multiple species coexist on the long term even in the limit $\nu \to 0$. Interestingly, the ecologically most relevant case, $D = 2$, is the critical dimension of this model. We shall see that this fact is a source of non-trivial behaviors of ecologically relevant quantities.

Duality is also an extremely powerful tool for computational analyses [21, 22]. If one is interested in the static, long-term, properties of the voter model with speciation, it is numerically much more efficient to simulate the dual dynamics than the forward one. In a dual simulation, after all walkers coalesced or were annihilated, species can be assigned to the start site of each walker, obtaining a stationary configuration of the voter model. Beside computational speed, this approach has also the advantage of eliminating finite-size effects induced by the boundary conditions, as the coalescing random walkers can be simulated in a virtually infinite system. For illustrative purposes, in Fig. 2c and 2d we show two configurations of the $2D$ voter model obtained with the dual dynamics for two different dispersal kernels.

C. $\beta-$diversity

The first ecological pattern we consider is the β-diversity, which is a measure of how the species composition in an ecosystem varies with the distance. We define the β-diversity as the probability $F(\vec{r})$, that two randomly chosen individuals at a distance $\vec{r}$ are conspecific, i.e. belong to the same species. We remark that, although this is the natural definition in this context, other definitions have been used in the ecological literature [40]. Mathematically, $F(\vec{r})$ can be expressed in terms of the two-point correlation function $G_{s_i,s_j}(\vec{r}) = \langle n_{s_i}(\vec{x}) n_{s_j}(\vec{x}+\vec{r})\rangle$, where $n_{s_i}(\vec{x})$ denotes the number of individuals of species s_i at location $\vec{x}$

$$F(\vec{r}) = \frac{\sum_i G_{s_i,s_i}(\vec{r})}{\sum_{i,j} G_{s_i,s_j}(\vec{r})}\,, \tag{1}$$

where the sums extend over all species in the ecosystem [11]. Eq. (1) can be used to estimate the β-diversity as the ratio between the couples of conspecific over the total number of couples in a sample.

Let us now study the evolution equation of $F(\vec{r},t)$ for the voter model with speciation and NN dispersal. Although we shall focus on the $2D$ case, it is useful to present the general calculation in D dimensions. Following [11, 41, 42] we write

$$\begin{aligned} F(\vec{r},t+1) \;&=\; \left(1-\frac{2}{N}\right) F(\vec{r},t) + \\ &+\; \frac{1-\nu}{DN}\sum_{k=1}^{D}\left[F(\vec{r}+\vec{e}_k,t)+F(\vec{r}-\vec{e}_k,t)\right]. \end{aligned} \tag{2}$$

The first term in the r.h.s. of Eq. (3) represents the fact that F does not change if two generic individuals at distance $\vec{r}$ are not removed in a given time step and therefore survive. The second term represents the events in which one of the two individuals dies (with prob. $2/N$), no speciation occurs (with prob. $1-\nu$) and the dead individual is replaced by a conspecific from the $2D$ neighbor

sites. Taking the continuous limit $N \to \infty$ with the lattice spacing $a \to 0$, the speciation probability $\nu \to 0$, and a finite value of $\kappa^2 = 2D\nu/a^2$, one obtains at stationarity the differential equation

$$\frac{1}{r^{D-1}}\frac{d}{dr}r^{D-1}\frac{dF}{dr} - \kappa^2 F(r) + c\delta^D(r) = 0 \qquad (3)$$

where δ^D is the $D-$dimensional Dirac delta, and because of isotropy the β-diversity $F(r)$ is now function of $r = |\vec{r}|$ only. The solution of Eq.(3) is [11]

$$F(r) = c\frac{\kappa^{D-2}}{(2\pi)^{D/2}}(\kappa r)^{(2-D)/2}K_{(2-D)/2}(\kappa r)\,, \qquad (4)$$

where K_z is the modified Bessel function of the second kind of order z and the constant c is fixed by the condition $\int_{r<a} d^D r F(\vec{r}) = 1$. We recall that Eq. (4) is a continuous expression, valid for distances much larger than the lattice spacing [41]. Although we derived Eq. (4) for NN dispersal, the same results hold for a general dispersal kernel for distances larger than the kernel range, provided that the kernel range is finite.

For $D = 2$, Eq. (4) implies that $F(r) \propto K_0(\kappa r)$, which is characterized by a slow logarithmic decay, $\sim -\ln(r\kappa)$, up to distances of order $1/\kappa \sim 1/\sqrt{\nu}$, followed by a faster, exponential falloff. Remarkably, the β-diversity empirically measured in several tropical forests in Central and South America is consistent with a logarithmic decay for large distances [43]. We remark that this logarithmic decay is the signature that $D = 2$ is the critical dimension for the voter model. In contrast, in $D = 1$, Eq. (4) becomes $F(r) \propto \sqrt{r\kappa}K_{1/2}(\kappa r) \sim \exp(-r\kappa)$. We mention for later convenience that, in $D = 1$ with NN dispersal, Eq. (3) can be solved without using the continuous approximation, giving [42]

$$F(r) = \exp(-\alpha(\nu)r)\,, \quad \text{with} \quad \alpha(\nu) = \ln\left[\frac{(1-\nu)}{(1-\sqrt{(\nu(2-\nu)})}\right]\,, \qquad (5)$$

where $\alpha(\nu) \approx \sqrt{(2\nu)}$ for $\nu \to 0$.

Although the β-diversity decays exponentially on scales $1/\kappa \sim 1/\sqrt{\nu}$ both in $1D$ and $2D$, there are important differences. Because $2D$ is the critical dimension, a large biodiversity (i.e. a large average number of species) can be sustained by very low values of the speciation rate ν. This implies that in $2D$ there are many species living on scales much smaller than $1/\kappa$, where the correlations decay logarithmically. Conversely, in $1D$ to maintain biodiversity one needs a large value of ν, so that $1/\kappa$ is the only characteristic scale and there is no additional structure on scales smaller than $1/\kappa$. This crucial point will be further elucidated in the rest of the section, where we will discuss other observables in $2D$ (subsections II D and II E) and compare them with their $1D$ counterparts.

D. Species-Area Relationships

We now focus on the SAR, defined as the average number of species, S of a given taxonomic level occupying a given area of size A. SARs are widely studied as a measure of spatial biodiversity and quantify how larger habitats support more species than smaller ones [3]. Empirical measures of SARs at multiple scales often reveal three different regimes [2, 3, 7]. At small areas, the number of species increases rather steeply, nearly linearly, with the sampled areas. A similar steep increase is observed at very large, continental scales. Instead, at intermediate scales, a slower, sublinear growth is often found. Such a growth is well approximated by a power law $S \sim A^z$, $z < 1$, over a wide range of taxa [44], though a logarithmic behavior $S \approx C \ln A$ has also been proposed. An extensive meta-study by Drakare et al. [45] reconsidered a large body of SAR studies from the literature, revealing that the power law provides a better fit in about half of the cases. This study also observed that the exponent z correlates positively with the body size of the considered group of species, so that small microorganisms typically display very shallow SAR curves as compared with larger organisms (see also [46] and Sect. III A).

Simulations of the (dual) voter model with speciation yields SARs qualitatively similar to those obtained from field data, see Fig. 3a. In the voter model, the steep initial regime is mostly determined by the dispersal range K. For areas significantly larger than K^2, a sublinear growth is observed (see Fig. 3b. In this regime, the growth becomes progressively more shallow as the speciation rate ν is decreased. For larger scales, the logarithmic slope of the SAR curves become steeper again. The area at which this final crossover occurs increases as ν is decreased.

An interesting question is whether the sublinear growth regime in the voter model can be characterized by a power-law $S \sim A^z$ and, in this case, what is the value of the exponent z as a function of ν. To address this question, we begin by reviewing a classic estimate of z by Durrett and Levin [20] relying on duality (see Sect. II B). The speciation rate ν sets a time scale $1/\nu$ which also corresponds to a characteristic length scale $\xi = 1/\sqrt{\nu}$ because of the diffusive behavior of random walkers in the dual model. Walkers with an initial separation much larger than ξ are likely to be annihilated before coalescence occurs. This observation alone explains the linear scaling of $S(A)$ for areas $A \gg \xi^2 = \nu^{-1}$. At these scales, species are uncorrelated, as can also be inferred from the analysis of the β-diversity in the previous section. For a system of coalescing random walkers in $2D$, the density of occupied sites $\rho(t)$ decays asymptotically as [47, 48]

$$\rho(t) \sim \frac{\ln t}{\pi t}\,. \qquad (6)$$

The characteristic logarithmic coarsening of clusters observed in the $2D$ voter model without speciation can be related to the logarithm appearing in Eq. (6) [49]. Assuming $\nu \ll 1$, the annihilation rate at time t in an area $\xi \times \xi$ can be approximated as the annihilation rate per walker ν times the number of walkers in the absence of annihilations $\xi^2\rho(t)$. Integrating over time, we find that

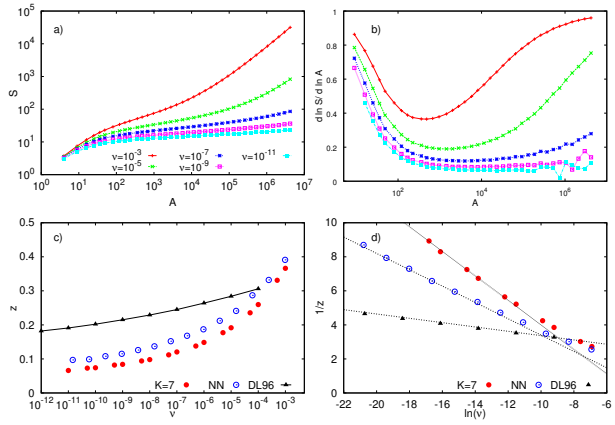


FIG. 3: Species Area Relationships (SAR) and their scaling behavior in the voter model with speciation. a) Number of species S as a function of the sampled area A for different speciation rates as in the caption. The triphasic shape is evident for larger speciations rates. Simulations from [22] were performed with a square dispersal kernel, i.e. $P(\vec{r})$ is a uniform distribution on a square of side K centered on the empty site, with $K = 7$. b) Local slopes, $d \ln S/d \ln A$ for the curves shown in panel a. c) Dependence of the exponent z on ν as obtained from the local slopes for both the square kernel with $K = 7$ and NN dispersal. The exponent is estimated at the inflection point of the SADs, i.e. at the minimum of the local slopes. Also shown is the prediction of Eq.(8) (black solid line) where the black triangles correspond to the values provided in [20]. d) Plot of $1/z$ vs $\ln(\nu)$ of the same data of panel c to highlight the logarithmic behavior of Eq. (9)

the total number of annihilations, i.e. the total number of species, is [39]

$$S(\xi^2) \approx \nu\xi^2 \int_{t_0}^{1/\nu=\xi^2} dt\ \rho(t) = \frac{\ln^2(\xi^2) - \ln^2(t_0)}{2\pi} \approx \approx \frac{2}{\pi} (\ln \xi)^2 , \quad (7)$$

where t_0 is the time at which the asymptotic expression (6) starts to be valid. The upper temporal cut-off is set to $1/\nu$ (with $1/\nu = \xi^2$) because the number of killing events occurring after a time $\sim 1/\nu$ is bounded by the number of walkers in the system, which is $\xi^2\rho(1/\nu) \sim \ln \xi$ [39]. Finally, combining Eq. (7), the fact that $S(1) = 1$ and matching a power law behavior $S = A^z$ in the range of scales from $A = 1$ to $A = \xi^2$, one finds [20]

$$z = \frac{\ln [S\,(A)]}{\ln\,(A)} = \frac{2\ln[\ln(1/\sqrt{\nu})] + \ln(2/\pi)}{\ln(1/\nu)} . \quad (8)$$

Also in this case, the logarithmic dependence of the exponent z on ν derives from the fact that $D = 2$ is the critical dimension for the voter model.

More recent results disputed the validity of Eq. (8). Scaling arguments hinted that z should approach a finite value $z \approx 0.2$ in the limit of vanishing ν (see [42] and Sec. II E 1), while numerical simulations suggested a power law dependence, $z \sim \nu^{0.15}$ [21]. Finally, further numerical simulations, based on the dual representation of the voter model with speciation (see Sect. II B) and spanning a very wide range of speciation rates from 10^{-3} to 10^{-11} confirmed the logarithmic behavior predicted by Eq. (8) [22]. The exponents measured in such simulations, shown in Fig. 3c, are well fitted by a phenomenological expression of the form

$$z = \frac{1}{q + m \ln(\nu)} \quad (9)$$

which is consistent with Eq. (8) up to order $\ln \ln \nu$, see also Fig. 3d. However, fitted values of the prefactors q and m are not consistent with Eq. (8). This discrepancy is probably due to pre-asymptotic effects as well as to the approximation of assuming a power-law range between $A = 1$ and $A = \ln(1/\nu)$.

Let us briefly discuss the role of the dispersal kernel. As illustrated in Figs. 3c and 3d, a comparison between NN dispersal and a square dispersal kernel of range $K = 7$ demonstrates that the exponent z depends to some extent on the dispersal kernel. However, numerical evidence [21, 22] suggests that when the dispersal kernel range is large enough (approximately $K \geq 5$) the exponents are very weakly dependent on K. Moreover, SARs obtained with different values of K can be rescaled onto a universal function of A and ν via the transformation $S = f(A, \nu, K) = K^\chi \phi(A/K^\chi, \nu)$ with a fitted value of $\chi \approx 1.97$. To the best of our knowledge, a formal derivation of this scaling law and of the exponent χ is currently an open problem.

The non-trivial area dependence of the SAR results is a special feature of the critical dimension $D = 2$. To highlight this point, we now discuss the $D = 1$ case as comparison. This case is also relevant to describe quasi one-dimensional ecosystems, such as river basins [50]. For simplicity, we limit ourselves to the case of NN dispersal.

To the best of our knowledge, also in $D = 1$, an exact expression for the average number of species, $S(L)$, in a segment of length L is unknown. Nevertheless, it is possible to provide a lower and upper bound for $S(L)$. In $D = 1$, the density of walkers behaves as $\rho(t) \sim 1/\sqrt{t}$, to be contrasted with eq. (6) valid in the $2D$ case. Dimensional arguments then suggest that the average number of species must a function of $L\sqrt{\nu}$ only, i.e. $S(L;\nu) = \Psi(L\sqrt{\nu})$. Computational results (Fig. 4a and inset) support well this simple argument. As shown in the figure, the non-trivial power-law regime characteristic of $2D$ SARs is absent in $D = 1$. Indeed, the function Ψ is linear for large arguments, with a coefficient around 1.2 and it is nearly constant for $L\sqrt{\nu} \ll 1$.

We can derive an upper bound to $S(L)$ using that, in $D = 1$, individuals are organized in $N_s(L;\nu)$ segments of conspecific individuals, so that $S \leq N_s$, with the equality holding if no species is present in more than one segment. We compute N_s from the probability $P_{i-1,i} \equiv F(|i - j|)$, with $F(r)$ given by Eq. (5), that two sites i and j are

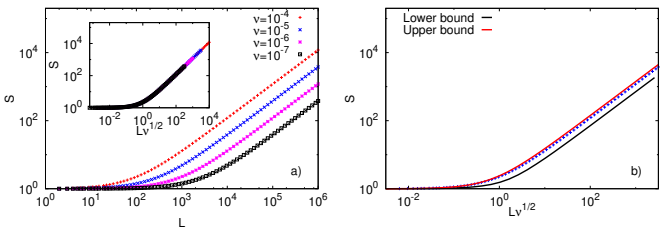


FIG. 4: Species Area Relationshipt for the voter model in $D = 1$. a) Average number of species S versus the system size L for different ν as labeled. Inset: same curves plotted vs $L\sqrt{\nu}$, notice the excellent collapse. b) SAR for $\nu = 10^{-5}$ compared with the theoretical upper (10) and lower (11) bounds.

occupied by conspecific individuals [51]

$$S \leq N_s = L - \sum_{i=1}^{L-1} P_{i-1,i} = L - (L-1)F(1) = \\ = L - (L-1)e^{-\alpha(\nu)}\,, \qquad (10)$$

which for $\nu \to 0$ can be approximated as $N_s \approx 1 + \sqrt{2\nu}(L-1)$.

The lower bound follows from Jensen's inequality (see also [51]) applied to the frequency of species represented by the individual in site $i \in [0, L-1]$, here denoted $n_i(L)$, which yields

$$S = \sum_i \left\langle \frac{1}{n_i(L)} \right\rangle \geq \sum_i \frac{1}{\langle n_i(L) \rangle}\,, \qquad (11)$$

where $\langle n_i \rangle = \sum_j P_{i,j}$ and $P_{i,j} = F(|i-j|)$ is again given by Eq. (5) and can be easily summed numerically.

In Fig. 4b we compare the numerically obtained SAR with the upper (10) and lower (11) bounds. Notice that the upper bound is very close to the actual SAR, implying that most species are organized in single segments.

E. Species-Abundance Distributions

We now discuss Species-Abundance Distributions (SADs), $P(n; A)$, that measure the relative abundance of species in a given area A. More precisely, denoting $S(A)$ the total number of species sampled in an area A, each composed by n_i $(i = 1, \ldots, S(A))$ individuals, $P(n; A)dn$ is the probability that a randomly picked species has an abundance between n and $n + dn$. While the expression of $P(n; A)$ for well-mixed neutral models is known [14], computing it for spatially explicit models, such as the voter model with speciation, has proven to be a rather hard problem. We first discuss in section II E 1 an approach based on standard finite-size scaling, and underline its limitations. In Sec. II E 2, we discuss how this approach can be generalized at the critical dimension, present numerical results, and discuss a recent attempt to compute $P(n; A)$ exploiting duality. Although we focus ond comparing the scaling theory with results from the voter model with speciation, we remark that the theoretical approach presented in this section is more general and can be applied to a vast class of models at the critical dimension.

1. *Power-law scaling relation*

In the voter model with speciation, the SAD is not only a function of the system size A, but also of the speciation rate ν. Although we are mainly interested in $2D$, it is instructive to consider the general case in which $A = L^D$, where L is the linear size of the sample. Following [11, 42], we assume a standard scaling form for the SAD

$$P(n; A, \nu) = n^{-\beta}\Psi(n\nu^{\alpha}, A\nu^{D/2}) \qquad (12)$$

where the exponents α and β remain unspecified for the time being, whereas the exponent $D/2$ stems from the diffusive nature of neutral models $\nu \sim t^{-1} \sim L^{-2} \sim A^{-2/D}$. Note that in models with long-range, non-diffusive dispersal [52] the scaling form might differ. Equation (12) describes a power-law dependence of P on n, holding up to a scale determined by the scaling function Ψ, that depends on dimensionless combinations of the population size n, the speciation rate ν, and the system size A. To the best of our knowledge, there is no available analytical prediction for the exponent β. The exponent α can be estimated in the dual formulation of the voter model with speciation, where the population size n is the number of coalescences that occur before an annihilation (see Sec. II B). This implies that α is the same exponent characterizing the temporal decay of the density of coalescing random walkers, $\rho(t) \sim t^{-\alpha}$. However, $\rho(t)$ decays as $\rho(t) \sim t^{-\min(1,D/2)}$ for $D \neq 2$ and $\rho(t) \sim \log(t)/t$ in $D = 2$, see eq. (6) and [48, 53]. Consequently, one should expect the power-law scaling of Eq. (12) to hold in $D = 1$ and $D \geq 3$, but not at the critical dimension $D = 2$, where logarithmic corrections should appear.

2. *Generalized scaling relation*

In order to allow for logarithmic corrections, Zillio *et al.* [30] proposed the generalized scaling relation

$$P(n; A) = g(A)\Psi(n/f(A))\,. \qquad (13)$$

The dependence on ν was omitted as the above scaling law was applied to observational data for which the speciation rate is unknown and assumed to be fixed. The key aspect of Eq. (13) is that f and g, are general functions and not necessarily power-laws as in conventional scaling, allowing for the possibility to include logarithms

or other functional dependencies. The scaling function $\Psi(x)$ is still assumed to be a power law

$$\Psi(x) \sim x^{-\Delta} \tag{14}$$

for small values of x, where Δ is an exponent to be determined. Thus, also Eq. (13) postulates a power-law dependence on n, but with a more general cut-off for large areas. After specifying the functions f and g, Eq. 13 can be tested by plotting $P(n;A)/g(A)$ versus $x = n/f(A)$ for a set of different areas and assessing the quality of the data collapse onto a single curve, $\Psi(x)$.

To determine the functions f and g, we impose that $P(n;A)$ has to be normalized, $\int_{n_0}^{\infty} dn\ g(A)\ \Psi(n/f(A)) = 1$, and that its average value has to be $\langle n\rangle = \int_1^{\infty} dn\ n\ g(A)\ \Psi(n/f(A))$. Substituting the scaling form (14) into these two equations, it is possible to derive conditions that the functions f and g must obey, depending on the value of Δ. In particular, the case $\Delta = 1$ is marginal and needs to be treated with care (other values $\Delta \neq 1$ are analyzed in the Appendix). Approaching such a limit as $\Delta = 1 - \epsilon$ with $\epsilon \ll 1$, Eq.(14) becomes

$$\Psi(x) = x^{-1+\epsilon} \sim \frac{1}{x}\big[\exp(\epsilon)\ln(x)\big] \sim \frac{1}{x}[1 + \epsilon \ln(x)] \tag{15}$$

up to first order in ϵ. At the same order in ϵ, the two conditions for $P(n;A)$ become $1 \sim g(A)f(A)\ln(f(A))[1 + \frac{\epsilon}{2}\ln(f(A))]$ and $\langle n\rangle \sim g(A)f(A)^2$, respectively, from which we finally obtain

$$\begin{aligned} f(A) &= \langle n\rangle \ln\langle n\rangle \left[1 + \frac{\epsilon}{2}\ln\langle n\rangle\right] \\ g(A) &= \frac{1}{\langle n\rangle \ln^2\langle n\rangle \left[1 + \frac{\epsilon}{2}\ln\langle n\rangle\right]^2} \end{aligned} \tag{16}$$

up to first order in ϵ. Notice that both functions f and g include logarithmic corrections. By means of a similar calculation, one can estimate the k-th moment $\langle n^k\rangle$, and verify that all the moment ratios $\langle n^k\rangle/\langle n^{k-1}\rangle$ scale in the same way, up to a multiplicative constant

$$\frac{\langle n^k\rangle}{\langle n^{k-1}\rangle} = \frac{\int dn\ n^k\ P(n;A)}{\int dn\ n^{k-1}\ P(n;A)} \propto f(A) \qquad k \geq 1\,. \tag{17}$$

revealing a highly anomalous scaling.

Zillio *et al.* [30], showed that this scaling form provides a much better collapse of empirical data from the Barro Colorado tropical forest than a power-law scaling relation such as Eq. (12). This supports the idea that Δ is close to its marginal value 1 in tropical forests.

We tested computationally whether Eqs. (13) and (16) provide a good collapse of SADs obtained from the voter model with speciation and whether the relationship between the moments, Eq.(17), holds. In simulations, an additional parameter is the speciation rate ν. As discussed above, ν appears in scaling relationships via the dimensionless combination $A\nu^{D/2}$, that in $2D$ equals $A\nu$. Thus, although Eqs. (13) and (16) do not include speciation explicitly, we expect these relationships to hold if

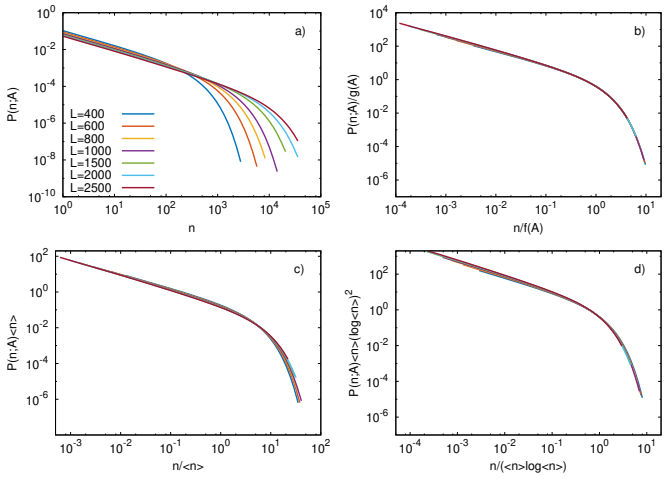


FIG. 5: SAD and data collapse. Results are presented for different linear system sizes and different speciation rates ν, keeping the product $A\nu = 200$ constant. a) SADs for different linear sizes from $L = 400$ to $L = 2500$. b) Collapse of SADs by means of Eqs.(13) and (16). The fitted parameter in the functions f and g is $\epsilon = 0.08$. c) Naive collapse without logarithmic corrections, where deviation for perfect collapse are evident. d) Collapse with the scaling form of Eqs.(13) and (16), but setting $\epsilon = 0$. Also in this case the discrepancy is evident.

$A\nu$ is kept constant. We therefore performed computational analyses fixing $A\nu = 200$, although the conclusions are robust against this choice. Results are summarized in Figure 5 which shows plots of the SAD, for systems with different linear size, L and different speciation rates ν (with $L^2\nu = A\nu = 200$). Observe in Fig. 5a that the smaller the size (or the larger the speciation rate) the smaller the maximal abundance. Figure 5b show the data collapse as given by Eqs. (13) and (16), where $\langle n\rangle$ is the average number of individuals measured in each area A and ϵ is a free parameter that we fitted obtaining $\epsilon = 0.08$ and a remarkable collapse of the different curves. The small value of ϵ, is consistent with the assumed small deviation from $\Delta = 1$. A similar collapse for $A\nu = 20$ leads to an even smaller value $\epsilon \approx 0.069$ (not shown). We verified that either removing all logarithmic corrections (thus plotting results as a function of $\langle n\rangle$) or simply fixing $\epsilon = 0$ in Eq. (13) and (16) leads to less convincing collapses, as shown in Fig. 5c and 5d, respectively. Clearly, these deviations can pass unnoticed in the presence of statistical fluctuations. Probably, this is the reason why in [54] a simple scaling law was claimed to hold for the $2D$ voter model with speciation. Finally, we also verified that moment ratios scale as $f(A)$, as predicted by Eq.(17) and illustrated in Fig.6.

In summary, a non-standard scaling form, including logarithmic corrections, provides an excellent collapse both for empirical data and for numerical simulations of the $2D$ voter model. We remark that the scaling theory is phenomenological, and the small parameter ϵ controlling the importance of logarithmic corrections is, at this

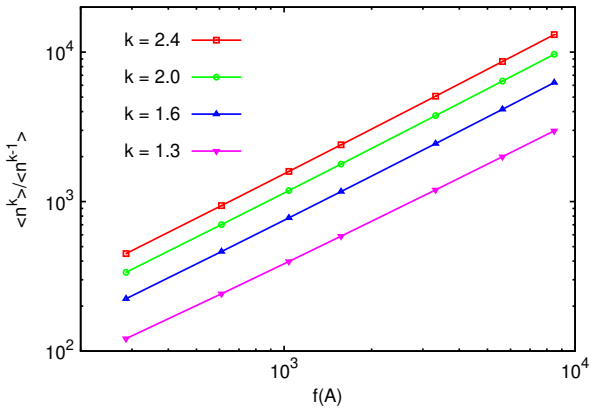


FIG. 6: Moment ratios for different values of k. As predicted by Eq.(17), in the case $\Delta \approx 1$ all moment ratios $\langle n^k\rangle/\langle n^{k-1}\rangle$ scale in the same way with $f(A)$ up to a multiplicative constant. As in Fig. 5, the fitted value is $\epsilon = 0.08$.

level, a non-universal free parameter. These results are in sharp contrast with the one-dimensional case, where logarithmic corrections are not expected. Indeed, Fig. (7) shows that the naive scaling form $P(n;A)\langle n\rangle$ vs. $n/\langle n\rangle$ (derived in Appendix A for the case $\Delta \neq 1$) yields a perfect collapse for SADs in one-dimensional systems.

It is interesting to remark that the data collapsed in [30] were obtained from tropical forests of different areas A. It is reasonable to assume that the speciation rate ν do not vary much among these forests. Therefore, the product $A\nu$ is not fixed, as in our computational analyses. A possible explanation is that, although the collapse achieved in this way is not perfect, the deviations from perfect scaling are too small to be appreciated in observational data due to the limited sample size. We have verified in simulations (not shown) that keeping ν constant (rather than $A\nu$ constant) small deviations from perfect collapse are observed.

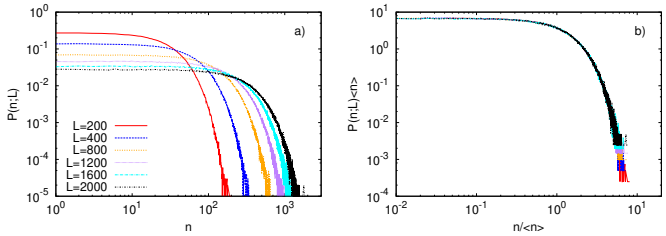


FIG. 7: Species Abundance Distribution (SAD) in the $D = 1$ voter model with speciation. a) SAD $P(n;L)$ vs n at varying the system size L as labelled with $A\nu = 40$ constant. b) Collapse of curved in (a) obtained withe the rescaling SAD $P(n;L)\langle n\rangle$ vs $n/\langle n\rangle$.

We conclude this section mentioning that a heuristic expression for the SAD has been recently derived for the voter model with speciation following a completely different approach [31, 32]. Let us define $P(x,t)$ as the distribution of the number of individual of a given species at time t. If we approximate x as a continuous quantity, we can heuristically write a Fokker-Planck equation for the evolution of $P(x,t)$

$$\partial_t P(x,t) = \nu\partial_x[xP(x,t)] + \partial_x^2[I(x)P(x,t)] \qquad (18)$$

where the first term in the right hand side is the negative drift due to speciation, and the second is the fluctuation in population size, where $I(x)$ is the average number of interfaces of a species of size x. The crucial underlying approximation is to neglect fluctuations of $I(x)$, which is appropriate if the distribution of the number of interfaces at fixed value of x is a very peaked function. In this simple framework, all the dependence on the spatial dimension of the voter model is recap into the function $I(x)$. The steady-state solution of Eq. (18) is

$$P_{st}(x) = \frac{e^{-\nu\int dx\frac{x}{I(x)}}}{I(x)}. \qquad (19)$$

From duality considerations [31, 32], the average number of interfaces must scale in $2D$ as $I(x) = x/(1 + c\ln x)$ where c is a non-universal constant. Notice how the expression of $I(x)$ includes familiar logarithmic terms and that the constant c plays the role of the exponent ϵ in the scaling theory. Substituting this expression into Eq. (19) leads to an explicit expression for the SAD, which obeys a scaling law with logarithmic corrections similar to Eq. (16), though not identical. A more detailed comparison between this result and the previous scaling form is an interesting issue, but beyond the scope of this review.

F. Species persistence-times

So far, we have considered neutral predictions of static ecological observables. However, neutral theory can also be used to predict time-dependent properties. A chief example is the distribution of survival times. The survival time τ (also called "persistence time") within a geographic region is defined as the time occurring between the speciation event originating a given species and its local extinction [15]. Recent empirical work on north-american birds and herbaceous plants revealed that the probability of observing a persistence time τ decays as as power laws $P(\tau) \sim \tau^{-1.83}$ and $P(\tau) \sim \tau^{-1.78}$ respectively, with area-dependent exponential cut-offs [55, 56].

In the voter model with speciation, the survival probability as a function of time can be computed analytically. Also in this case, the calculation relies on duality [47, 48, 57]. In $2D$ and in the limit of vanishing ν one obtains

$$P(\tau) \sim \frac{\ln\tau}{\tau^2} \qquad (20)$$

while standard power-law scaling $P(\tau) \sim \tau^{-1/2}$ is expected in 1D. For non-negligible values of ν, these scaling forms are cut-off by a ν-dependent exponential factor $\exp(-\nu\tau)$ in either dimension. Also in this case, diffusive scaling relates the characteristic time scale $1/\nu$ with a length scale ξ via $\xi \sim \sqrt{\nu}$. This explains the aforementioned area-dependent cut-offs observed in empirical data [55].

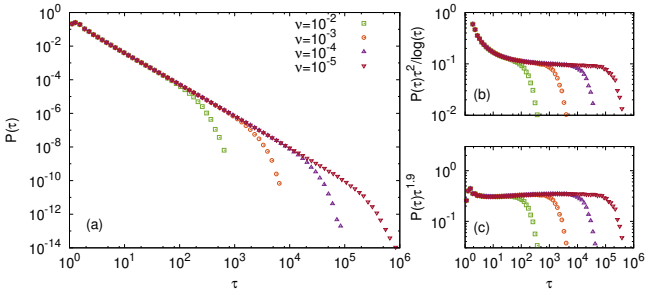


FIG. 8: Species persistence times. (a) Probability distribution function of species persistence times for different values of the speciation rate ν as in label. (b) and (c) show the pdf rescaled with the logarithmic correction, $P(\tau)\tau^2/\ln\tau$, and with a power law, $P(\tau)\tau^{1.9}$, respectively.

Species persistence times in simulations of the $2D$ voter model with speciation are shown in Fig. 8a. Panels (b) and (c) show compensated plots of the simulation results. The simulations support the prediction of eq. (20) (panel b), and also illustrate that a power law with an exponent close to 2 (1.9 in this case) provides a good approximation of the scaling predicted by Eq. (20) in a broad range of scales (panel c), consistently with the empirical findings in [55, 56].

III. OTHER NEUTRAL MODELS

In the voter model with speciation, the habitat is saturated and each site is always occupied by an individual. In this section, we study neutral spatial models where the number of individuals that can inhabit a site is varied. We consider three variants: the stepping-stone model with speciation, where each site can host many individuals but the landscape remains saturated; the contact process with speciation, where occupancy is limited to a maximum of one individual per site, but sites can also be empty; and the O'Dwyer-Green model, where occupancy is unbounded.

A. Stepping-Stone Model with speciation

In the voter model, each lattice site hosts a single individual. This assumption is appropriate for big sessile species, such as trees, where each individual occupies a well-defined area and exploits its local resources. On the other side of the spectrum, microorganisms, such as small eukaryotes or bacteria, are often present in very large numbers on tiny spatial scales, where all individuals share the same resources. For these species, it is more appropriate to think of the habitat as subdivided into small patches, connected by migration and each hosting a large number of individuals directly competing with each other [58]. To model such ecological cases, in this section we consider the stepping-stone model [25, 26] with speciation, which generalizes the voter model with speciation to the case in which each site hosts a fixed number M of individuals.

Similar to the voter model with speciation, at each time step an individual is randomly chosen and killed. With probability ν, it is replaced by an individual of a novel species. With complementary probability $(1-\nu)$, a reproduction event occurs. The parent of the new individual is selected with probability $(1-\mu)$ among the surviving $M-1$ individuals present at the same site, and with probability μ among the M individuals in a randomly chosen neighboring patch (according to a probability distribution on the neighbors $P(\vec{r})$, similar to the case of the voter model). The particular case of $M=1$ reduces to the voter model with speciation up to a time rescaling $t \to \mu t$. Like the voter model, the stepping-stone model admits a dual representation in terms of coalescing random walkers with annihilation, which can be exploited for efficient numerical simulations. The main difference with respect to the dual of the voter model is that, in the dual stepping-stone model, at each step a random walker can either move to another site or stay in the site of origin. Coalescence can happen in both circumstances, corresponding to reproduction of an individual from neighboring sites or from the same site. For full details on the implementation we refer to [27].

As revealed by numerical simulations of the stepping-stone model based on the dual representation, SARs are qualitatively similar to those of the voter model, although the exponents z are, in general, smaller than in the voter model [27]. In particular, the exponent depends not only on ν, but also on the combination of parameters $M\mu$, which determines the regimes of the model. For $M\mu \ll 1$, each local site is likely to contain only one species. In this limit, each site behaves as one individual up to a time rescaling, so that one should expect the same exponents as in the voter model with speciation. In the opposite limit $M\mu \gg 1$, there is a large diversity of species at each site. An analytical argument suggests that, in this latter limit, the exponent should be a factor two smaller than in the former limit [27]. Let us study the limit $M\mu \gg 1$ in the dual representation. Since random walkers in the same site have a low probability of coalescence, they will wander for a long time before coalescing. Therefore, we can assume that, asymptotically, they will behave as in the well-mixed case. This implies that their density in an area smaller or equal than ξ^2 approximately decays

according to the mean-field formula

$$\rho(t) \sim \frac{1}{t} \ . \tag{21}$$

Observe that in this case the characteristic length is $\xi = \sqrt{\mu/\nu}$, as random walks diffuse with probability μ at each time step. Proceeding as in Eq. (7), the average number of species in an area ξ^2 can be estimated as

$$S(\xi^2) \sim \nu M \xi^2 \int_{t_0}^{\mu/\nu=\xi^2} \frac{dt}{t} = M\mu \ln\left[\frac{\xi^2}{t_0}\right] . \tag{22}$$

To compute z, we also need an estimate for $S(1)$, that in this case is not trivially equal to one. As the population is assumed to be well-mixed in an area equal to ξ^2 or smaller, the composition of a single site can be thought as a sample of M individuals from this well-mixed population. The probability distribution of the abundance in such a sample is given by Ewens' sampling formula [59]. Substituting its expression yields

$$S(1) = \sum_{j=0}^{M-1} \frac{M\mu}{M\mu + j} \approx M\mu \ln(1+\mu^{-1}) \ . \tag{23}$$

Combining Eqs. (22) and (23) and assuming a power law in the range from $A = 1$ to $A = \xi^2$, we find an exponent

$$z \sim \frac{\ln(\xi^2)}{\ln\ln(\xi^2)} = \frac{\ln\ln(\nu/\mu)}{\ln(\nu/\mu)} \tag{24}$$

which, to the leading order, is a factor 2 smaller than the corresponding estimate for the voter model (8). The decrease of the exponent z with the combination of parameters $M\mu$ is confirmed in numerical simulation, see Fig. 9, although the asymptotic reduction is less than the factor two predicted by the approximate estimate of eq. (24).

Summarizing, the stepping-stone model at large local community size M yields smaller values of the species-area exponent z than the voter model [27]. This fact is consistent with the ecological observation that microbial communities, characterized by very large local community sizes, typically display very shallow species-area relations, and that in general there seems to be a positive correlation between the exponent z and the body size of a taxonomic group [46]. In the stepping-stone model, a decrease in the SAR exponent is observed in the regime $M\mu \gg 1$ where each site hosts a large number of species and therefore provides a buffer for biodiversity [27]. This interpretation is also consistent with the "cosmopolitan" nature of many microbial species, i.e. the fact that relatively small communities of microbes host a biological diversity comparable with that observed in the whole planet [58, 60]. This feature has sometimes been explained invoking the fact that microbes have the possibility of long-range dispersal [60]. However, numerical simulations show that, in the voter-model with speciation, long-range dispersal leads to steeper, rather than shallower SARs [52].

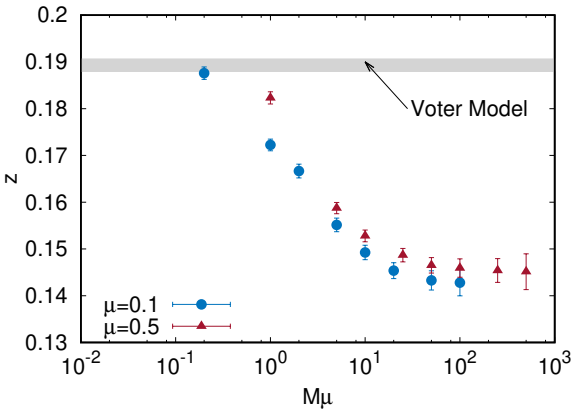


FIG. 9: Species-area exponents for the Stepping Stone Model at fixed $\nu = 10^{-6}$, different local population size M and dispersal rate μ, with NN dispersal. The numerical estimate of the exponent z in the voter model for NN-dispersal and the same value of the speciation rate is shown for comparison.

B. Contact Process with speciation

In the voter model, every dead individual is instantly replaced by a newborn, leading to a constantly saturated environment. The implicit underlying assumption is that the birth rate is infinite, so that death events are the rate-limiting steps. Such assumption constitutes a good approximation in resource-rich ecosystems. In less rich ecosystems, where the birth rate is finite, the environment is not always saturated and empty gaps can exist [61].

To explore this latter case, we study here the contact process with speciation, which is the multi-species variant of the well-known contact process [18, 28, 29, 38, 62]. As usual, we consider the model on a $2D$ square lattice. Sites of the lattice can be occupied by individuals belonging to different species or empty. The model is defined in continuous time; each individual dies at a rate d and reproduces at a rate b. In case of a death, the site is simply left vacant. A reproduction event is considered successful only if the individual has at least one vacant neighboring site. In such a case, one of the vacant neighboring sites is chosen at random. With probability ν, the site is occupied by an individual of a new species (speciation event); with complementary probability, $(1-\nu)$, the newborn is of the same species as the parent.

As in the standard contact process [38, 62], the parameter determining the asymptotic density of occupied sites ρ is the dimensionless birth-to-death ratio $\eta = b/d$. For $\eta < \eta_c \approx 1.649$ the absorbing state in which all sites are empty is stable. A non-equilibrium phase transition at $\eta = \eta_c$ separates this region from a stable active phase ($\eta > \eta_c$) characterized by a non-vanishing value of ρ that depends on η [28, 29]. For $\eta \to \infty$ one has $\rho \to 1$ and the model is equivalent to the voter model with speciation [18].

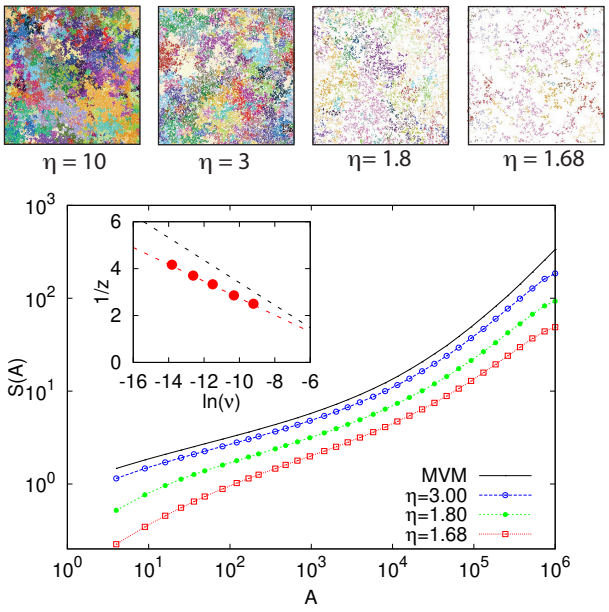


FIG. 10: (top) Snapshots of configurations of the contact process with speciation at different values of the birth-to-death rate ratio η and $\nu = 10^{-4}$. In each panel, each color represents a different species. (bottom) SARs at different values of the birth-to-death rate η (shown in the figure legend) and $\nu = 10^{-5}$. (inset) Red dots: estimated exponent z as a function of ν for the contact process with speciation at $\eta = 1.68$. Red dashed line is a linear fit; black dashed line is the corresponding fit for the voter model with speciation for comparison. We have chosen a NN dispersal kernel in all panels.

The CP is a self-dual model. Therefore, duality cannot be exploited in numerical simulations as in the case of the voter model. Forward simulations show that the SAR and the corresponding exponents are remarkably similar to the voter model with speciation even at small values of η, corresponding to very fragmented ecosystems as shown in Fig. 10. For values of η very close to η_c (but within the active phase) and small values of ν, SAR exponents tend to be smaller than in the voter model, see inset of Fig. 10.

In principle, in a very fragmented ecosystem it would not make sense to sample empty areas, or areas with very few individuals. With this idea in mind, an alternative to the standard definition of SAR used so far is to weigh the sample of a given area with its number of individuals, i.e. of occupied sites. Adopting this definition one finds qualitatively different SARs for small values of η [27]. In particular, these SARs do not seem to be characterized by a clear power-law range. We refer the reader to Ref. [27] for a broader discussion of this issue.

C. O'Dwyer-Green model

We have seen that finding exact results for neutral spatial models constitutes a formidable problem, and even in the simple case of the voter model only asymptotic results are known.

To make progress in this direction, O'Dwyer and Green proposed a spatial neutral model in which individuals do not compete, i.e. the site occupancy is not bounded [63]. In their model, each individual can reproduce at a rate b, giving rise to a newborn located according to a dispersal distribution, die at a rate d, or speciate at a rate ν, giving rise to a newborn of a new species. The model was studied at the critical point $b+\nu = d$. The lack of interaction considerably simplifies the mathematical treatment: the model can be mapped into a field theory from which the authors of [63] obtained an analytical expression for the species-area law and the dependence of z on ν. In particular, the solution was derived by writing an equation for the distribution of a generic species, which was solved by imposing detailed balance. However, Grilli and coworkers [64] pointed out a flaw in this procedure. In this model all species are transient, as the birth rate of each species is always smaller than the death rate because of speciation. This implies that all species eventually go extinct, so that the detailed balance (i.e. equilibrium) assumption is not valid.

An often overlooked aspect of the O'Dwyer and Green model is the lack of a carrying capacity. Although well-mixed neutral models commonly do not have a carrying capacity (beside that of the entire ecosystem), a local carrying capacity, i.e. a maximum occupancy of each lattice site, is a standard ingredient in spatial neutral theory, shared by all models we discussed so far. In the O'Dwyer and Green model, since the dynamics of the entire ecosystem is a critical branching process, the population at each site undergoes huge fluctuations. This fact implies as a drawback that numerically simulating the steady-state of the model and sampling its configurations is extremely difficult. While the authors of [64] clearly pointed out that the detailed balance solution leads to several inconsistencies and is therefore not valid, to the best of our knowledge there have been no attempt of comparing this solution with numerical simulations to see if detailed balance can provide a reasonable approximation of the dynamics in some particular regimes or limits.

Currently, the research of spatial neutral models that can be solved analytically is still open [65]. In this direction, although this review focuses on lattice models, we mention a recent phenomenological attempt based on a spatial Fokker-Planck equation where both space and population sizes are continuous variables [66].

IV. NEAR-NEUTRAL MODELS

In the previous sections, we focused on neutral ecological models. However, in real ecosystems the neutral

assumption is (at best) a crude approximation. It is thus interesting to examine some of the main effects of non-neutral forces, also because many biodiversity patterns that are well predicted by neutral models are also found in richer, non-neutral models [67–69]. A main difficulty in comparing neutral and non-neutral models is the large number of possible ecological effects (and corresponding parameters) that typically enter the latter. In this section, with the aim of understanding basic non-neutral effects in a simple setting, we present a minimal model introduced in [70], where one can continuously move from a neutral to a non-neutral scenario by varying a single parameter, tuning the amount of spatial disorder. We then discuss generalizations to other types of spatio-temporal disorder.

A. Habitat-preference model

We consider a variant of the voter model where different sites are preferred habitats for each one of the competing species. For the sake of simplicity, we limit ourselves to the case of two species A and B with N_A and N_B individuals, respectively. We assume habitat saturation, so that the total population is $N = N_A + N_B = L^2$ where the system is a square lattice of size L with periodic boundary conditions. Individuals of type A and B can also migrate to the system from an infinite reservoir where they are equally represented. Each lattice site can be of type a or b, i.e. being a preferred habitat for colonization by species A or B, respectively. After colonization, mortality and dispersal do not depend on being on a preference site. Ecologically, this means that the fitness advantage belongs to the seeds and not to the individuals themselves (see [71] for a different choice). The a vs b character of each site is chosen randomly at the beginning and it remains fixed over time – quenched disorder. To maintain the model globally symmetric, we assume equal proportions of a and b sites and that intensity of the two biases (a favoring A and b favoring B) are identical. The dynamics proceeds as follows. At each discrete time step, a lattice site is randomly chosen with uniform probability and the residing individual is killed. The individual is replaced either by an immigrant from the reservoir (with probability μ) or by an offspring of an individual residing in one of the four neighboring sites (with probability $(1-\mu)$). In both cases, the colonization probability is biased by an additional factor γ for the individuals that have preference for the empty site. In formulas, the probability of colonization of a site $x = \{a, b\}$ by an individual $X = \{A, B\}$ ($Y = \{B, A\}$) having (not having) preference for that site is

$$\begin{aligned} W_X^x(n_X, n_Y) &= (1-\mu)\tfrac{(1+\gamma)n_X}{(1+\gamma)n_X+n_Y} + \mu\tfrac{1+\gamma}{2+\gamma} \\ W_Y^x(n_X, n_Y) &= (1-\mu)\tfrac{n_Y}{(1+\gamma)n_X+n_Y} + \mu\tfrac{1}{2+\gamma}\,, \end{aligned} \tag{25}$$

respectively, where n_X (n_Y) denotes the number of individuals of species $X = \{A, B\}$ ($Y = \{B, A\}$) in the neighborhood of the considered site. Similar models have been proposed also in the context of heterogeneous catalysis [72] and social dynamics [73]. For $\gamma = 0$ and $\mu = 0$, the standard (neutral) voter model with two species is recovered. For $\gamma = 0$ but $\mu \neq 0$, it corresponds to the noisy voter model [74, 75].

Also in this model, the results can depend on the choice of the dispersal kernel $P(\mathbf{r})$. Here we focus on the NN dispersal and global dispersal (GD), i.e. a mean-field version of (25). The GD case can be thought as a variant of the two islands model [76] of population genetics, where each island host $N/2$ individuals and is favorable to one of the two species. In the mean-field version, the state of the system is univocally determined by the numbers of individuals N_{Aa} and N_{Bb} residing on their island of preference. The numbers of individuals outside their island of preference are $N_{Ba} = N/2 - N_{Aa}$ and $N_{Ab} = N/2 - N_{Bb}$. The dynamics is then fully specified by the probabilities per elementary steps that N_{Xx} (with $X = \{A, B\}$ and $x = \{a, b\}$) increases or decreases by a unit:

$$\begin{aligned} \mathcal{W}_{N_{Xx}\to N_{Xx}+1} &= \left(\frac{1}{2} - \frac{N_{Xx}}{N}\right) W_X^x(N_A, N_B) \\ \mathcal{W}_{N_{Xx}\to N_{Xx}-1} &= \frac{N_{Xx}}{N}\, W_Y^x(N_A, N_B) \end{aligned} \tag{26}$$

where W_Y^x and W_X^x are given by eqs. (25) with n_X and n_Y replaced by $N_X = N_{Xx} + N_{Xy}$ and $N_Y = N_{Yy} + N_{Yx}$, respectively.

B. Extinction times

In the absence of immigration ($\mu = 0$) and for finite populations $N < \infty$, persistent coexistence of the two species is not possible: demographic stochasticity eventually drives one of the species to extinction (the absorbing state) with the *fixation* (in the jargon of population genetics) of the other species. In this case, information on the system can be obtained by studying the dynamics toward extinction [71]. Of particular interest is the average extinction time, $\langle T_{ext}\rangle$, and its dependence on system properties, such as the deviation from neutrality and the population size.

In the neutral case ($\gamma = 0$), as discussed, the system recovers the voter model with NN dispersal and the Moran model [37] in the version with global dispersal. In this limit, the extinction time is set by the population size. In particular, for large N we have $\langle T_{ext}\rangle \sim N \ln N$ for NN-dispersal [77] and $\langle T_{ext}\rangle \sim N$ for global dispersal [37, 78]. To inquire the effect of habitat preferences we performed simulations of the model (25) with an initial condition $N_A = N_B = N/2$ until the extinction of one of the two species.

Figure 11 shows the average extinction time, measured in generations, i.e. N elementary steps of eqs. (25), as a function of the population size N for different values of γ. For $\gamma = 0$ we observe the $N \ln N$ behavior expected in

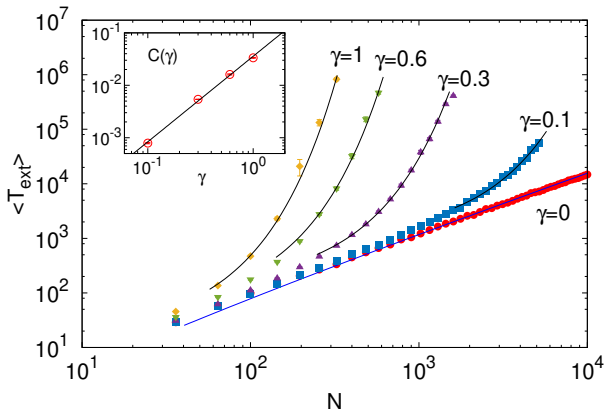


FIG. 11: Extinction times for the model with NN dispersal without immigration ($\nu = 0$). Mean extinction time $\langle T_{ext}\rangle$ as a function of N for different values of γ as in label. The blue curve approximating the neutral $\gamma = 0$ data points corresponds to the neutral expectation $\langle T_{ext}\rangle \propto N \ln N$, the black curves over the symbols for $\gamma \neq 0$ correspond to exponential fits of the form $\langle T_{ext}\rangle \propto \exp(C(\gamma)N)$. The inset shows (symbols) $C(\gamma)$ vs γ, while the black solid line display the best fit $C(\gamma) = A\gamma^{\beta}$ with $\beta \approx 1.63$. The average extinction time is obtained by an annealed average, i.e. by randomizing the preference sites at each realization. Each point represents an average over 10^3 realizations.

the neutral case. Habitat preference ($\gamma > 0$) leads to a dramatic increase of the average extinction time, which becomes exponential in N

$$\langle T_{ext}\rangle \propto \exp(C(\gamma)N)\,, \tag{27}$$

for large enough N. The dependence of the constant $C(\gamma)$ on γ, shown in the inset, is well-fitted by a power-law with exponent ≈ 1.63. The mean-field version of the model presents similar qualitative features with the only difference that $\langle T_{ext}\rangle \propto N$ for $\gamma = 0$ and with some differences in the γ dependence of $C(\gamma)$, as shown in [70].

The exponential dependence of the average extinction times on N indicates that habitat preference has a stabilizing impact on the population dynamics. Indeed, when N is large enough, the two species coexist on any realistic time scale. The stabilizing effect of habitat preference reflects also in the probability of fixation P_{fix}, i.e. the probability that a species, say A, gets fixated when initially present as a fraction $x = N_A/N$ of the population. In the neutral case, standard computation [78] shows that $P_{fix}(x) = x$. As shown in [70], when γ is increased, $P_{fix}(x)$ develops a much steeper dependence on x and quickly reaches values $\approx 1/2$ even for small x, provided that γ is large enough. In other words, the stabilization due to habitat preference tends to compensate any initial disproportion between the population of the two species.

C. Coexistence

In the presence of immigration ($\mu > 0$), a locally extinct species can recolonize, leading to a dynamical coexistence between the two species. However, if the typical recolonization time $1/\mu$ is large compared to the average extinction time $\langle T_{ext}\rangle$, such recovery from extinction is slow and unlikely. Therefore, most of the time the ecosystem is dominated by one of the two species. Therefore, the distribution of the population size of any of the two species, $P(X)$ ($X = A, B$) is peaked at 0 and at the population size N, corresponding to dominance of either of the two species. We denote this regime as *monodominance*, see Fig. 12a. In the opposite limit $\langle T_{ext}\rangle \gg 1/\mu$, temporary extinctions are very unlikely and the distribution is peaked at $N_A = N_B = N/2$ leading to *pure coexistence* of the two species (Fig. 12c). For intermediate values of μ, temporary extinctions are still possible though the replenishment due to immigration will tend to equilibrate the two populations. In this case of *mixed coexistence*, the distribution is characterized by three local maxima at $N_X = 0, N/2, N$ (Fig. 12b).

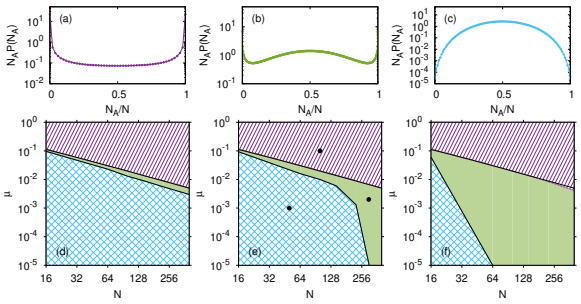


FIG. 12: Different regimes of coexistence for the case with NN dispersal and immigration for the model with habitat preference. Top panels show the stationary distribution $P(N_A)$ for $\gamma = 0.3$ and (a) $N = 50$ with $\mu = 10^{-3}$, (b) $N = 300$ with $\mu = 2 \times 10^{-3}$, and (c) $N = 100$ with $\mu = 10^{-3}$, corresponding to a typical distribution in the cases of monodominance, mixed regime and pure coexistence, see text. Bottom panels show how the three regimes partition the N, μ-parameter space for different values of γ: (d) $\gamma = 0$ corresponding to the neutral case, (e) $\gamma = 0.3$ and (f) $\gamma = 1$. The three points in (e) correspond to the distributions displayed in the top panels, as labelled by the color coding.

Figs. 12d,e,f show the three regimes of coexistence in the $N - \mu$ parameter space for the model with NN-dispersal for different habitat preference strength γ (increasing from left to right). In the mean-field model, we find the same qualitative features, except that for $\gamma = 0$ the mixed regime is absent, so that one has a direct transition from monodominance to pure coexistence [70].

The main emerging feature is that increasing habitat preference expands the region of parameter space corresponding to mixed coexistence at the expenses of monodominance. Surprisingly, the pure coexistence regime

seems to be insensitive to the degree of habitat preference. In particular, the critical line $\mu_c(N)$ separating it from the mixed regime seems to be the same that separates coexistence from monodominance in the neutral model ($\gamma = 0$) with global dispersal, which is given by the expression $\mu_c(N) = 2/(2+N)$. This result can be obtained in the following way. For $\gamma = 0$, the transition rates (26) can be expressed in terms of the rates for N_A to increase/decrease by one

$$\mathcal{W}_{N_A \to N_A \pm 1} = \frac{\frac{N}{2} \pm \left(\frac{N}{2} - N_A\right)}{N} \times \times \left[(1-\mu)\frac{\frac{N}{2} \mp \left(\frac{N}{2} - N_A\right)}{N} + \frac{\mu}{2}\right]. \quad (28)$$

Then, the equilibrium distribution $P(N_A)$ can be computed imposing the detailed-balance condition

$$\frac{P(N_A+1)}{P(N_A)} = \frac{\mathcal{W}_{N_A \to N_A+1}}{\mathcal{W}_{N_A \to N_A-1}}, \quad (29)$$

which must hold at stationarity since the process is one dimensional [79]. To determine $\mu_c(N)$ for the transition from monodominance to coexistence, it is sufficient to determine whether, for small N_A, $P(N_A)$ is an increasing or a decreasing function. Using (29) with (28) and imposing $P(N_A+1) > P(N_A)$ one obtains, after some algebra, the inequality $[(2+N)\mu - 2](N - 2N_A - 1) > 0$, which is verified whenever $\mu > 2/(2+N)$. Notice that, in the case of global dispersal, the distribution is uniform along this line, i.e. for $\mu = \mu_c$ one finds $P(N_A) = 1/N$.

D. Generalizations of the habitat-preference model

To gain physical insight into the different regimes shown in Fig. 12, a variant of the habitat preference model was introduced and analyzed for the global dispersal case in [80]. By considering the first two terms of a system-size expansion of the master equation, results in the infinite-size limit and finite-size corrections were derived. In the infinite-size limit, i.e. neglecting the effect of fluctuations, the introduction of a non-vanishing local preference generates a deterministic force, which can be described as an effective potential $V(\delta)$ for the relative difference of densities $\delta = (N_A - N_B)/N$. This potential has a minimum at the coexistence state, $\delta = 0$, corresponding to a maximum in the probability distribution at $N_A = N_B = N/2$. In other words, species coexistence emerges for infinitely large sizes. On the other hand, for finite systems, when fluctuations are considered, the only possible true steady states are the absorbing states $\delta = \pm 1$, where the effective potential $V(\delta)$ is singular. The minimum at δ is separated from the negative singularities by two potential barriers. As strength of the local preference and/or N increase, the basin of attraction of the coexistence state becomes larger and deeper and the two symmetric barriers become closer to the absorbing states and higher. Consequently the time needed to escape the coexistence state becomes much longer, therefore unaccessible in computer simulations. Thus, three different regimes can thus be identified: the absorbing, intermediate (quasi-active) and active phase (much as in Fig. 12). In the absorbing phase, symmetry is broken and one of the two species reaches extinction with certainty. This regime is equivalent to the monodominance regime in Fig. 12. The active phase is characterized by a coexistence of both species, and survives fluctuations only in the infinite-size limit. This corresponds to the coexistence phase of Fig. 12. Finally, the intermediate state is a mixture of the two previous ones: the absorbing states and the coexistence state are locally stable, thus, the system is tri-stable, and the steady state depends on the initial conditions. This is the mixed state of Fig. 12. These results provide a nice analytical example of how noise can effectively change the shape of a deterministic potential. Still, the presence of absorbing states – with the associated singularities in the steady state distribution – prevent true phase transitions from occurring: the only possible steady state for any finite system is an absorbing one. Only in the infinite-size limit, noise vanishes and the coexistence state becomes truly stable [80].

Another study scrutinized the case in which there are local habitat preference only at some specific locations in space, while all other sites are neutral [81]. An interesting example which has been analyzed in details is that of a square lattice where only the left (resp. right) boundary has a preference for species A (resp. B), ([81], see also [82, 83]). The conclusion is that even mild biases at a small fraction of locations induce robust and durable species coexistence, also in regions arbitrarily far apart from the biased locations. As carefully discussed in [81] this result stems from the long-range nature of the underlying critical bulk dynamics of the neutral voter model, and is robust to the introduction of non-symmetrical biases –i.e. stronger for one of the species– except for the fact that the state of coexistence is no longer symmetric. These conclusions have a number of potentially important consequences, for example, in conservation ecology as it suggests that constructing local "sanctuaries" for different competing species can result in global increase of stability of their populations, and thus an enhancement of biodiversity, even in regions arbitrarily distant from the protected zones [81].

E. Temporally-dependent habitat preferences

We have seen that spatial quenched disorder generically fosters species coexistence. Another important question is what happens when the preference for a species are time-dependent, i.e. if neutrality is temporarily broken in favor of one of the coexisting species, while the ecosystem remains neutral on average. This question has a long tradition in ecology. Several theoretical studies have looked at the impact of environmental fluc-

tuations on population growth and ecosystem stability [71, 84]. On one hand, environmental stochasticity enhances fluctuations and extinction rates, that can have a destabilizing effect on the ecological community. On the other hand, it can also foster stability, as the temporal alternance of species can effectively reduce the strength of interspecific competition.

Similarly to the case of spatial disorder, one can design quasi-neutral models where habitat-preferences for different species are time-dependent, i.e. where in each time window there is a preference for a randomly chosen species. Different works have recently analyzed this type of models, showing that time-dependent habitat preference greatly improves predictions of empirical ecological patterns with respect to purely neutral theories [31, 85–87]. In particular, it has been claimed that these models provides more realistic estimates of dynamical quantities, such as average species persistence times and distributions of species turnover [88], compared with their neutral counterparts.

F. Models with density dependence

In ecology, one speaks of *density-dependence* or Allee effect when the fitness of an individual depends on the abundance of the species it belongs to. The underlying mechanisms can be very diverse, from cooperative defense/feeding to spreading of parasites among conspecific. An interesting scenario is that of negative density-dependence, i.e. when individuals belonging to more abundant species have lower fitness. It is established that, in well mixed systems, negative density-dependence significantly favors species coexistence [89]. Versions of the voter model implementing a negative density-dependence have been studied in the literature [90, 91]. In these models, the reproduction probability of an individual depends on the number of conspecific individuals in a given local neighborhood. Strictly speaking, these models are not neutral: the neutral hypothesis is defined at the level of individuals [7], and here individuals belonging to species of different abundance do not have the same fitness. However, these models, as the other models considered in this Section, are still symmetric, since all species are treated on equal footing. Interesting phenomena like the possibility of spontaneous breakdown of such a symmetry –thus leading to asymmetric species coexistence– have been recently uncovered at the mean field level [92].

V. PERSPECTIVES AND CONCLUSIONS

The range of ecological problems discussed in this review is by no means exhaustive, and we believe there are many directions that still need to be explored or fully understood.

A prominent example is the role of different speciation mechanisms on spatial biodiversity. In the models discussed in this review, speciation events involve a single individual (*point speciation mode*, in the language of evolutionary ecology). This assumption is convenient from the modeling perspective, but leads to fitted values of the speciation rate that are incompatible with independent estimates [93]. This assumption also tends to generate too many young species which last for a short time and overweights rare species. To address these issues, recently, another mechanism called *protracted speciation* has been proposed in the context of neutral models [94]. In protracted speciation, the speciation event does not occur at a single generation, but is a gradual event lasting for some generations. Introducing protracted speciation partially solves some of the aforementioned problems [94]. In real ecosystems, even more speciation mechanisms are at play [95]. For example, in *parapatric speciation*, two spatially-separated population of the same species can diverge and give rise to two different species. This would correspond to a speciation event involving a group of individuals rather than a single one. The role of different speciation modes in maintaining biodiversity and in patterning the spatial organization of species is still under discussion and modeling results can provide very useful contributions to this debate.

As mentioned in the Introduction, ecological neutral theory elicited a heated debate which is far from being solved as, in many cases, non-neutral models based on the concept of niche and neutral models yield similar fits of biodiversity patterns [67, 68, 96]. In recent years a new view on this debate has been emerging. In Chase and Leibold's words: "niche and neutral models are in reality two ends of a continuum with the truth most likely in the middle" [97]. Indeed, the ecological forces underlying niche and neutral models are not mutually exclusive, and demographic stochasticity plays an important role also in non-neutral settings. However, it has been difficult to clarify the importance of different neutral and non-neutral mechanisms, as most non-neutral model are characterized by a large number of parameters. Some progress in this direction has been obtained in simplified settings which, similarly to the model presented in Sect. IV, allow for a controlled departure from neutrality. For instance, Haegeman and Loreau [98] added the main ingredients of neutral theory, demographic stochasticity and immigration, to a Lotka-Volterra competition model. Similar problems have been studied in Refs. [99–101]. An interesting future direction would be to study similar models in a spatial context.

In many ecological communities, in particular of microbial organisms, ecological and evolutionary timescales are not separated. Eco-evolutionary models describing both processes are becoming more and more important [102]. Neutral theory has provided a simple framework to describe patterns in these communities, for example in gut microbiota [103]. These systems call for new theoretical efforts and new observables, such as generalizations

of the β-diversity taking into account genetic differences among individuals [104].

We have seen throughout this review how some observables measured by ecologists corresponds to well known quantities in statistical physics: for instance, the β-diversity is closely related with a two-point correlation function. Other observables, such as SARs and SADs, are less common in statistical physics. A potentially fruitful future direction is to consider other observables which are common in statistical mechanics, such as multi-point correlation functions, and measure them in ecosystems. In this direction, it is very interesting the study of species clustering in [105] based on the theory of continuum percolation.

In summary, we presented an overview of different stochastic spatial models in population ecology. We have seen that even very simple models are a source of challenging problems in statistical physics. In particular, because of speciation, each species is bound to extinction and is therefore ultimately transient. This feature is in contrast with traditional classical spin system defined on a lattice where, even when in out-of-equilibrium conditions, the number of spin components is fixed from the beginning. Further, ecosystems are typically two-dimensional and, due to the underlying diffusive behavior, $D = 2$ is the critical dimension for these models. We have shown that this fact often leads to logarithmic corrections to scaling laws, which have been difficult to analyze both analytically and numerically. Despite these difficulties, remarkable progress has been made in recent years. We believe that cross-fertilization between statistical physics and ecology will be more and more important in the future to deepen our quantitative understanding of how ecosystems are organized.

Appendix: General scaling relationships

In this brief Appendix, we discuss general condition imposed on the functions f and g by the properties of the function Ψ, depending on the exponent Δ, see eq.(13), eq. (14) and [30]. Let us write the normalization condition for $P(n; A)$

$$\sum_n P(n; A) \approx g(A)f(A) \int_{n_0/f(A)}^{\Lambda} dx\ x^{-\Delta} = 1. \quad (30)$$

The infrared cutoff Λ is related to the fact that the function $\psi(x)$ is a power-law for small x only and rapidly decays for larger arguments, see e.g. Fig. 5. The integral is singular for small x and $\Delta > 1$ and thus

$$1 \sim g(A)f(A)f(A)^{\Delta-1} = g(A)f(A)^{\Delta}\ . \quad (31)$$

On the other hand, if $\Delta < 1$, the integral is weakly dependent on $f(A)$, so that

$$1 \sim g(A)f(A)\ . \quad (32)$$

Similarly, the first moment of Ψ is

$$\langle n\rangle \sim g(A)f^2(A)f(A)^{\Delta-1} = g(A)f(A)^{\Delta+1} \quad (33)$$

if $1 < \Delta < 2$ and

$$\langle n\rangle \sim g(A)f^2(A) \quad (34)$$

for $\Delta > 2$. Combining the expressions above, different regimes emerge as a function of Δ: if $\Delta < 1$, $f(A) = \langle n\rangle$, while for $1 < \Delta < 2$, $f(A) = \langle n\rangle^{1/(2-\Delta)}$, while no specific prediction for $f(A)$ can be made in the case $\Delta \geq 2$. In particular, for $\Delta < 1$ one has a simple scaling form $f(A) = \langle n\rangle$ and $g(A) = 1/\langle n\rangle$ which applies, for example, to the $1D$ case as described in the main text. The marginal case $\Delta = 1$ is treated in detail in Sec. II E.

Acknowledgments

MAM is grateful to the Spanish-MINECO for financial support (under grant FIS2013-43201-P; FEDER funds), as well as to J. Hidalgo, S. Suweis, A. Maritan, C. Borile for a long term collaboration on topics related to the content of this paper.

[1] M. L. Cody and J. M. Diamond, *Ecology and evolution of communities* (Harvard University Press, 1975).
[2] F. Preston, Ecology **41**, 611 (1960), ISSN 0012-9658.
[3] M. Rosenzweig, *Species diversity in space and time* (Cambridge Univ Pr, 1995), ISBN 0521499526.
[4] R. MacArthur, Am. Nat. **94**, 25 (1960).
[5] M. Tokeshi, Adv. Ecol. Res. **24**, 111 (1993).
[6] S. A. Levin, Ecology **73**, 1943 (1992).
[7] S. Hubbell, *The unified neutral theory of biodiversity and biogeography Princeton University Press* (Princeton, New Jersey, USA, 2001).
[8] M. Kimura, *The neutral theory of molecular evolution* (Cambridge University Press, 1983).
[9] D. Alonso, R. S. Etienne, and A. J. McKane, Trends Ecol. Evol. **21**, 451 (2006).
[10] J. Rosindell, S. P. Hubbell, and R. S. Etienne, Trends Ecol. Evol. **26**, 340 (2011).
[11] S. Azaele, S. Suweis, J. Grilli, I. Volkov, J. R. Banavar, and A. Maritan, Rev. Mod. Phys. **88**, 035003 (2016).
[12] M. Vallade and B. Houchmandzadeh, Phys. Rev. E **68**, 061902 (2003).
[13] A. J. McKane, D. Alonso, and R. V. Solé, Theor. Pop. Biol. **65**, 67 (2004).
[14] I. Volkov, J. Banavar, S. Hubbell, and A. Maritan, Nature **424**, 1035 (2003).
[15] S. Pigolotti, A. Flammini, M. Marsili, and A. Maritan,

Proc. Nat. Acad. Sci. **102**, 15747 (2005).
[16] R. S. Etienne and D. Alonso, J. Stat. Phys. **128**, 485 (2007).
[17] J. Rosindell, S. P. Hubbell, F. He, L. J. Harmon, and R. S. Etienne, Trends Ecol. Evol. **27**, 203 (2012).
[18] R. Durrett and S. A. Levin, Philos. Trans. Royal Soc. London B: Biol. Sci. **343**, 329 (1994).
[19] R. S. Etienne and J. Rosindell, PloS one **6**, e14717 (2011).
[20] R. Durrett and S. Levin, J. Theor. Biol. **179**, 119 (1996).
[21] J. Rosindell and S. Cornell, Ecol. Lett. **10**, 586 (2007), ISSN 1461-0248.
[22] S. Pigolotti and M. Cencini, J. Theor. Biol. **260**, 83 (2009), ISSN 0022-5193.
[23] T. Liggett, *Interacting particle systems* (Springer Verlag, 1985), ISBN 3540226176.
[24] M. Kimura, Ann. Rept. Nat. Inst. Genetics, Japan **3**, 62 (1953).
[25] M. Kimura and G. H. Weiss, Genetics **49**, 561 (1964).
[26] K. S. Korolev, M. Avlund, O. Hallatschek, and D. R. Nelson, Rev. Mod. Phys. **82**, 1691 (2010).
[27] M. Cencini, S. Pigolotti, and M. A. Muñoz, PloS one **7**, e38232 (2012).
[28] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).
[29] H. Hinrichsen, Adv. in Phys. **49**, 815 (2000).
[30] T. Zillio, J. R. Banavar, J. L. Green, J. Harte, and A. Maritan, Proc. Nat. Acad. Sci. **105**, 18714 (2008).
[31] Y. Shem-Tov, M. Danino, and N. M. Shnerb, Sci. Rep. **7** (2017).
[32] M. Danino, Y. Shem-Tov, and N. M. Shnerb, arXiv preprint arXiv:1606.02837 (2016).
[33] P. Clifford and A. Sudbury, Biometrika **60**, 581 (1973).
[34] O. A. Pinto and M. A. Munoz, PloS one **6**, e21946 (2011).
[35] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[36] W. Croft, Selection **3**, 75 (2002).
[37] P. A. P. Moran, in *Math. Proc. Camb. Philos. Soc.* (Cambridge University Press, 1958), vol. 54, pp. 60–71.
[38] R. Holley and T. Liggett, Ann. Prob. **3**, 643 (1975), ISSN 0091-1798.
[39] M. Bramson, J. Cox, and R. Durrett, Ann. Prob. **24**, 1727 (1996).
[40] H. Tuomisto, Ecography **33**, 2 (2010).
[41] J. Chave and E. G. Leigh, Theoretical population biology **62**, 153 (2002).
[42] T. Zillio, I. Volkov, J. R. Banavar, S. P. Hubbell, and A. Maritan, Phys. Rev. Lett. **95**, 098101 (2005).
[43] R. Condit, N. Pitman, E. G. Leigh, J. Chave, J. Terborgh, R. B. Foster, P. Núnez, S. Aguilar, R. Valencia, G. Villa, et al., Science **295**, 666 (2002).
[44] O. Arrhenius, J. Ecol. **9**, 95 (1921), ISSN 0022-0477.
[45] S. Drakare, J. J. Lennon, and H. Hillebrand, Ecol. Lett. **9**, 215 (2006).
[46] M. Horner-Devine, M. Lage, J. Hughes, and B. Bohannan, Nature **432**, 750 (2004), ISSN 0028-0836.
[47] M. Bramson and J. L. Lebowitz, J. Stat. Phys. **62**, 297 (1991).
[48] L. Peliti, J. Phys. A: Math. Gen. **19**, L365 (1986).
[49] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, Phys. Rev. Lett. **87**, 045701 (2001).
[50] S. Suweis, E. Bertuzzo, L. Mari, I. Rodriguez-Iturbe, A. Maritan, and A. Rinaldo, J. of Theor. Biol. **303**, 15 (2012).
[51] B. Derrida and B. Jung-Muller, J. Stat. Phys. **94**, 277 (1999).
[52] J. Rosindell and S. Cornell, Ecology **90**, 1743 (2009), ISSN 0012-9658.
[53] M. Bramson and J. Lebowitz, Phys. Rev. Lett. **61**, 2397 (1988).
[54] J. Rosindell and S. J. Cornell, Oikos **122**, 1101 (2013).
[55] E. Bertuzzo, S. Suweis, L. Mari, A. Maritan, I. Rodríguez-Iturbe, and A. Rinaldo, Proc. Natl. Acad. Sci. **108**, 4346 (2011).
[56] S. Suweis, E. Bertuzzo, L. Mari, I. Rodriguez-Iturbe, A. Maritan, and A. Rinaldo, J. Theor. Biol. **303**, 15 (2012).
[57] B. P. Lee, J Phys. A: Math. Gen. **27**, 2633 (1994).
[58] T. Fenchel and B. J. Finlay, BioScience **54**, 777 (2004).
[59] W. Ewens, Theor. Pop. Biol. **3**, 87 (1972).
[60] B. Finlay and T. Fenchel, Protist **155**, 237 (2004), ISSN 1434-4610.
[61] M. Loreau, Ecol. Lett. **3**, 73 (2000).
[62] D. Griffeath, Stoch. Proc. Appl. **11**, 151 (1981).
[63] J. P. O'Dwyer and J. L. Green, Ecol. Lett. **13**, 87 (2010).
[64] J. Grilli, S. Azaele, J. R. Banavar, and A. Maritan, EPL (Europhysics Letters) **100**, 38002 (2012).
[65] J. P. O'Dwyer and S. J. Cornell, arXiv preprint arXiv:1705.07856 (2017).
[66] S. Azaele and F. Peruzzo, bioRxiv p. 074336 (2016).
[67] B. McGill, Oikos **102**, 679 (2003).
[68] D. Tilman, Proc. Nat. Acad. Sci. **101**, 10854 (2004).
[69] B. Gilbert and M. J. Lechowicz, Proc. Nat. Acad. Sci. **101**, 7651 (2004).
[70] S. Pigolotti and M. Cencini, J. Theor. Biol. **265**, 609 (2010).
[71] P. L. Chesson and R. R. Warner, Am. Nat. **117**, 923 (1981).
[72] L. Frachebourg, P. Krapivsky, and S. Redner, Phys. Rev. Lett. **75**, 2891 (1995).
[73] N. Masuda, N. Gibert, and S. Redner, Phys. Rev. E **82**, 010103 (2010).
[74] A. Kirman, Quart. J. Econom. **108**, 137 (1993).
[75] B. L. Granovsky and N. Madras, Stoch. Proc. Appl. **55**, 23 (1995).
[76] P. A. P. Moran, *The statistical process of evolutionary theory* (Clarendon Press, 1962).
[77] P. Krapivsky, Phys. Rev. A **45**, 1067 (1992).
[78] J. Gillespie, *Population genetics: a concise guide* (Johns Hopkins University Press, 2004), ISBN 0801880092.
[79] C. Gardiner, *Stochastic methods*, Springer Series in Synergetics (Springer-Verlag, Berlin, 2009), 2009).
[80] C. Borile, A. Maritan, and M. A. Muñoz, J. Stat. Mech.: Th. Exp. **2013**, P04032 (2013).
[81] C. Borile, D. Molina-Garcia, A. Maritan, and M. A. Muñoz, J. Stat. Mech.: Th. Exp. **2015**, P01030 (2015).
[82] M. Mobilia, Phys. Rev. Lett. **91**, 028701 (2003).
[83] M. Mobilia, A. Petersen, and S. Redner, J. Stat. Mech.: Th. Exp. **2007**, P08029 (2007).
[84] L. Ridolfi, P. D'Odorico, and F. Laio, *Noise-induced phenomena in the environmental sciences* (Cambridge University Press, 2011).
[85] M. Danino, N. M. Shnerb, S. Azaele, W. E. Kunin, and D. A. Kessler, J. Theor. Biol. **409**, 155 (2016).
[86] J. Hidalgo, S. Suweis, and A. Maritan, J. Theor. Biol. **413**, 1 (2017).

[87] T. Spanio, J. Hidalgo, and M. A. Muñoz, Phys. Rev. E **96**, 042301 (2017).
[88] S. Azaele, S. Pigolotti, J. R. Banavar, and A. Maritan, Nature **444**, 926 (2006).
[89] P. Chesson, Annual review of Ecology and Systematics pp. 343–366 (2000).
[90] J. Molofsky, R. Durrett, J. Dushoff, D. Griffeath, and S. Levin, Theoretical Population Biology **55**, 270 (1999).
[91] F. Schweitzer and L. Behera, The European Physical Journal B-Condensed Matter and Complex Systems **67**, 301 (2009).
[92] C. Borile, M. A. Muñoz, S. Azaele, J. R. Banavar, and A. Maritan, Phys. Rev. Lett. **109**, 038102 (2012).
[93] R. E. Ricklefs, Ecology **87**, 1424 (2006).
[94] J. Rosindell, S. J. Cornell, S. P. Hubbell, and R. S. Etienne, Ecology Letters **13**, 716 (2010).
[95] J. A. Coyne and H. A. Orr, *Speciation* (Sinauer Associates, Inc, 2004).
[96] J. Chave, H. C. Muller-Landau, and S. A. Levin, Am. Nat. **159**, 1 (2002).
[97] J. M. Chase and M. A. Leibold, *Ecological niches: linking classical and contemporary approaches* (University of Chicago Press, 2003).
[98] B. Haegeman and M. Loreau, J. Theor. Biol. **269**, 150 (2011).
[99] A. E. Noble, A. Hastings, and W. F. Fagan, Phys. Rev. Lett. **107**, 228101 (2011).
[100] A. E. Noble and W. F. Fagan, arXiv preprint arXiv:1102.0052 (2011).
[101] S. Pigolotti and M. Cencini, J. Theor. Biol. **338**, 1 (2013).
[102] P. Villa Martin, J. Hidalgo, R. Rubio de Casas, and M. Muñoz, Plos Comp. Biol **12**, e1005139 (2016).
[103] P. Jeraldo, M. Sipos, N. Chia, J. M. Brulc, A. S. Dhillon, M. E. Konkel, C. L. Larson, K. E. Nelson, A. Qu, L. B. Schook, et al., Proceedings of the National Academy of Sciences **109**, 9692 (2012).
[104] B. Houchmandzadeh, Physical Review E **95**, 012402 (2017).
[105] J. B. Plotkin, J. Chave, and P. S. Ashton, Am. Nat. **160**, 629 (2002).

Acknowledgements

> You cannot insult a man more atrociously than by refusing to believe he is suffering.
>
> Cesare Pavese, This Business of Living

Next, I thank those who have had an special relevance for this thesis. I will write it in Spanish since it is my mother language and I feel it closer to me.

Gracias a BCAM, por hacer posible mi trabajo durante estos años y tenerme en consideración. Me siento afortunado por haber entrado al mundo laboral de la mano de esta institución.

Gracias a mi supervisor de tesis, Gianni Pagnini, por colaborar en la creación de esta tesis y ofrecer su conocimiento y experiencia a la investigación subyacente.

Gracias a Paolo Paradisi por señalarme las técnicas de computación necesarias para la investigación.

Gracias a Ralf Metzler y a Alexei Chechkin por incluirme en su trabajo de investigación y enriquecer mi universo científico.

Gracias a Arghir Zarnescu, por su voluntad de mediar y hacer posible la escritura de esta tesis.

Gracias a todos aquellos de miembros de BCAM que alguna vez han tenido a bien hacerme la vida un poco más fácil, cada uno a su manera y dentro de sus posibilidades. Gracias de corazón.

Gracias a Kike, por ser un buen amigo y buen compañero de trabajo. No es fácil encontrar personas afines en las que poder confiar. Le deseo mucha suerte en su carrera profesional.

Gracias a mis padres, por muchas cosas que uno no suele pararse a pensar pero cumplen un papel tan esencial como desapercibido. En este periodo tengo presente una. A veces cuesta pedir ayuda. A mí me cuesta mucho. Tengo suerte de tener unos padres generosos en ofrecerla.